

Prediction of Finger Movements using Ultrasound Images

Emanuel Zarka



DIPLOMARBEIT

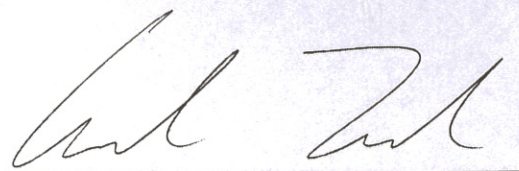
PREDICTION OF FINGER MOVEMENTS USING ULTRASOUND IMAGES

Freigabe:

Der Bearbeiter:

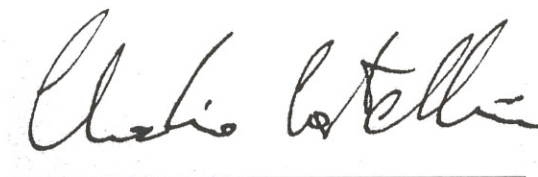
Unterschriften

Emanuel Zarka



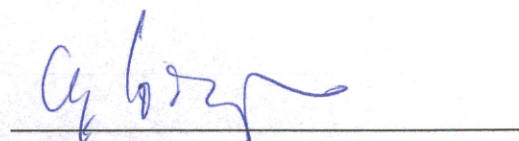
Betreuer:

Claudio Castellini



Der Institutsdirektor

Prof. Dr. G. Hirzinger



Dieser Bericht enthält 101 Seiten, 55 Abbildungen und 16 Tabellen

MASTER THESIS

Thesis submitted in partial fulfilment of the requirements for the degree of
Master of Science in Engineering

at the University of Applied Sciences Technikum Wien
Mechatronik / Robotik (Master)

Prediction of Finger Movements using Ultrasound Images

by
Emanuel Richard Zarka, BSc.
1210 Vienna, 55/41 Gerasdorferstrasse

Supervisors:
Claudio Castellini, Ph.D.
Peter Goldmann, Dipl. Ing.

Vienna, October 9, 2011

Declaration

I confirm that this paper is entirely my own work. All sources and quotations have been fully acknowledged in the appropriate places with adequate footnotes and citations. Quotations have been properly acknowledged and marked with appropriate punctuation. The works consulted are listed in the bibliography. This thesis has not been submitted to another examination panel in the same or a similar form, and has not been published.

Place, Date

Signature

Kurzfassung

Sonographie ist ein weit verbreitetes, bildgebendes Verfahren das neben der Pränataldiagnostik eine Vielzahl an weiteren Anwendungsmöglichkeiten findet. Mit Hilfe der unschädlichen, hochfrequenten Schallwellen können innere Organe, Muskeln, Sehnen und Knochen dargestellt werden. Die in dieser Arbeit vorgestellte Methode ermöglicht eine Rekonstruktion der Fingerpositionen mittels Ultraschallbildern des Unterarmes. Lokale Merkmale erster Ordnung werden aus den Graustufenbildern extrahiert und gleichzeitig die Gelenkwinkel der Finger mit einem Sensorhandschuh erfasst. Basierend auf diesen Daten wird eine Regressionsanalyse durchgeführt. Erste Versuche offenbaren einen linearen Zusammenhang zwischen den Bildmerkmalen und Fingerpositionen. Dies ermöglicht die Datenverarbeitung und Vorhersage der Fingerstellungen in Echtzeit. Ein auf dem Optical-Flow basierender Algorithmus gleicht etwaige, ungewollte Verschiebung der Ultraschallsonde aus. Das hier vorgestellte System kann unter anderem auf Grund des hohen Auflösungsvermögens zur Steuerung von Handprothesen und Roboterhänden eingesetzt werden. Experimente eines handamputierten Menschen untermauern die Einsatzmöglichkeit zur Phantomschmerztherapie.

Schlagwörter: Sonographie, Bildverarbeitung, Maschinelles Lernen, Prothetik

Abstract

Ultrasound imaging as a diagnostic device is a widespread, non-invasive technique used to investigate the interior of the human body. It uses high-frequency sound waves to visualise organs as well as musculoskeletal structures. In this thesis, ultrasound images of the human forearm are used to predict finger positions. A possible application for the presented method is the control of dexterous hand prostheses or robotic hands. Local first-order features extracted from ultrasound images, and position values gathered by a sensor glove, provide the basis for regression analysis. Preliminary experiments prove the feasibility of this approach and reveal that there is a linear relationship between image features and finger joint angles. This holds uniformly for several subjects and enables on-line computation. Specific finger movements, such as finger flexion, thumb rotation and thumb adduction, can be reconstructed up to sub-degree precision. In order for the system to correctly estimate the hand posture, movements of the ultrasound transducer with respect to the skin must be kept to a minimum. An optical flow based technique is implemented to compensate the inevitable movements and prevent drift errors. Several experiments conducted with an amputee also suggest that the system is able to reconstruct the imaginary limb and therefore may be used to treat phantom limb pain.

Keywords: Ultrasound imaging, image processing, machine learning, prosthetics

Acknowledgements

This thesis was conducted at the DLR's Institute of Robotics and Mechatronics in Oberpfaffenhofen. I would like to express my appreciation to my supervisor Dr. Claudio Castellini for his guidance, encouragement and support. My sincere thanks also goes to Dr. Georg Passig, Dr. Patrick van der Smagt and everyone from the DLR's Bionics Group. Last but not the least, I would like to thank Dipl. Ing. Peter Goldmann of the University of Applied Sciences Technikum Wien.

Contents

1	Introduction	1
2	State of the Art	3
3	Problem and Solution Approach	5
4	Theoretical Foundation	6
4.1	Ultrasonography	6
4.1.1	Requirements	6
4.1.2	Technical Principle	6
4.1.3	Static Imaging Principles	13
4.1.4	Real-Time Ultrasound Imaging Principles	14
4.1.5	Risks	16
4.2	Anatomy of the Human Arm	16
4.2.1	Terms of Direction and Terms of Movements	16
4.2.2	Muscles of the Human forearm and Hand	18
4.2.3	Bones of the Human Hand	23
4.3	Machine Learning and Pattern Recognition	24
4.3.1	Terminology	24
4.3.2	Linear Models for Regression	26
4.3.3	Support Vector Machine	30
4.4	Image Processing	30
4.4.1	Feature-specific Measurements	31
4.4.2	Gray Value Surfaces	31
4.4.3	Optical Flow	32
4.5	Phantom Limb Pain	33
4.5.1	Treatment of Phantom Limb Pain	33
5	Setup Specification	35
5.1	Used Hardware	35
5.1.1	Ultrasound Imaging System	36
5.1.2	VGA Framegrabber	39
5.1.3	Cyberglove	42
5.2	Used Software	43
5.2.1	Virtual Hand Model	44
5.3	Principle Procedure	44
5.3.1	Data Gathering	48
5.3.2	Training Phase	53
5.3.3	Testing Phase	56
5.4	Advanced Modifications	57

5.4.1	Probe Motion Compensation	57
5.4.2	Signal Filtering and Data Processing	60
6	Experiments Description	62
6.1	Preliminary Experiments	62
6.2	Final Experiments	65
6.2.1	Experiment 1	65
6.2.2	Experiment 2	65
7	Results and Discussions	67
7.1	Preliminary Results	67
7.1.1	Correlation Coefficient	72
7.2	Final Results	72
7.2.1	Questionnaires	72
7.2.2	Experiment 1	78
7.2.3	Experiment 2	80
7.3	Feature Properties	82
7.4	Discussions	83
8	Conclusion and Future Work	86
	List of Figures	89
	List of Tables	90
	List of Abbreviations	91
	Bibliography	92

1 Introduction

Today, medical robots fulfil a growing number of health related tasks for assisting individuals with disabilities. This includes devices for rehabilitation therapy, medical training or prosthetics. The goal is to provide greater independence to people with disabilities by helping them perform activities of daily living. For example, robot manipulators can assist individuals who have impaired arm or hand function with basic tasks such as eating and drinking. Robotic prostheses have been developed to replace missing parts of the human body and are used to provide mobility or manipulation abilities when a limb is lost (Comprehensive references to rehabilitation and medical robotics are provided by Hillmann (2004), and Speich and Rosen (2004)).

This thesis presents a non-invasive approach to predict the position of the hand. This opens up the possibility for below-elbow amputees to control dexterous robotic hands or polyarticular prostheses.

Upper extremity amputation can be related to accidental trauma, disease, defect or the progress of war. Although persons have suffered arm amputation for centuries, it is only during the past 500 years that substitutes have been fabricated in an attempt to provide a substitute for the amputated part of the arm. Much of the early writing that exists in reference to amputation, is related to war injuries. One of the first examples of an early artificial hand is a mailed fist that was made for Goetz von Berlichingen in the middle ages. This prosthesis was equipped with jointed fingers that could passively grips his sword like a vice. Even with the advances made in the electrical control of prosthetic part, man has not build an excellent substitute for a missing hand, wrist, elbow or shoulder (Meier and Atkins, 2004).

In the context of non-invasive interfaces for controlling mechanical hands, a concrete possibility arises from forearm surface electromyography (EMG), a technique by which muscle activation potentials are gathered by electrodes placed on the patient's forearm. These potentials can be used to track which muscles the patient is activating, and with what force (Castellini and van der Smagt, 2009). The use of myoelectric signals for control seems to have been suggested first by Reiter, in Germany, around 1948 (Wirta, et al., 1977).

In this thesis, however, another non-invasive technique is presented in which a medical ultrasound system gathers ultrasound images of the human forearm. The general idea of the feasibility of this approach stems from plain observation of the ultrasound imaging of the human forearm. In order to visualise the muscles and tendons, the ultrasound transducer is placed on the palmar side of the forearm. The ultrasound system generates 2D gray-scale images in which tissue interfaces are represented as bright ridges. As the subject's fingers are moving, a clear relationship between changes in the image and the hand configuration is apparent. Image features are used to extract the information.

Ultrasound as a clinical modality began in the mid-1950s for examination of the eye and for

echoencephalography. In the early to mid-1970s, ultrasound instrumentation changed significantly with the advent of analog and digital scan converters, which made two-dimensional imaging possible. The development of ultrasound instrumentation continued with the introduction of real-time scanners in the late 1970s, followed by Doppler and colour Doppler units in the 1980s (Hedrick, 1995).

The technique is easy to set-up and has no apparent side effects. One of its best known applications is the imaging of the foetus with pre-birth diagnostic purposes. Although ultrasound machines are not cheap, they are nowadays practically found in any hospital, making the technique easily accessible.

The original idea of this thesis was to use ultrasound images together with EMG signals. However, the very first experimental tests suggested that ultrasound imaging itself contains enough information to build a properly working system.

This thesis is organised as follows: after reviewing the related work, the problem and solution approach is shown. The theoretical foundations are explained before setup specifications and experiments are described in detail. Results are then presented and discussed, and finally, conclusion and future work is presented.

2 State of the Art

This thesis was conducted at the DLR's Institute of Robotics and Mechatronics (Deutsches Zentrum für Luft- und Raumfahrt, German Aerospace Center in Oberpfaffenhofen, Germany). Research carried out in this institute is based on the interdisciplinary design, computer-aided optimisation and simulation, as well as implementation of complex mechatronic systems and man-machine interfaces. The institute is regarded as one of the worldwide leading institutions in the field of robotics.

One of the most famous institute's robotic system is the humanoid robot called "Rollin Justin" with its compliant controlled light weight arms. It has two four-finger hands and is an ideal experimental platform for dual handed mobile manipulation. Various sensors allow the 3D reconstruction of the robot's environment and therefore enable Justin to perform given tasks autonomously (see Fuchs, et al., 2009). Other major developments of the institute concern hand arm systems and robotic hands. The DLR hand-arm-system consists of 52 drives, more than 100 sensors and mimics the kinematic, dynamic and force properties of the human arm. The DLR hands are more dexterous than other mechatronic hands, while maintaining excellent force-to-weight ratios (see for example Grebenstein, et al., 2011, and Liu, et al., 2008).

In order to improve robots and robotic systems, the DLR's Bionics Group studies human systems. Research topics include artificial skin, kinematics of the human hand, brain-controlled robotics and non-invasive rehabilitation and prostheses (see van der Smagt, et al., 2009a). Of particular interest is the use of surface EMG to detect the intention and force of movement of the patient, either amputated or left with lesser muscle functionality (van der Smagt, et al., 2009b).

Pattern recognition and machine learning techniques are commonly used as a intelligent base frame. For example, a method based on support vector machines which can detect opening and closing actions of the human fingers recorded via surface EMG. The method is successfully demonstrated on a robotic four-finger hand, and can be used to grasp objects (Bitzer and van der Smagt, 2006).

Also the work by Castellini and van der Smagt (2009) deals with advanced robotic hand control via EMG. It shows how machine learning, together with a simple downsampling algorithm can be effectively used to control both on-line and in real time, finger position as well as finger force of a highly dexterous robotic hand. The system determines the type of grasp a human subject is willing to use, and the required amount of force involved, with a high degree of accuracy. This represents a remarkable improvement with respect to the state-of-the-art of feed-forward control of dexterous mechanical hands, and opens up a scenario in which amputees will be able to control hand prostheses in a much finer way than it has so far been possible. The practicability of pattern recognition in the myoelectric control of multiple-axis upper-limb prostheses has been effectively demonstrated by Wirta, et al. (1977).

Ultrasound imaging is successfully used as a diagnostic tool for hand musculoskeletal disorders, for example rheumatoid arthritis (see De Flaviis, et al., 1988). This suggests that ultrasound images contain enough information to reconstruct the position, velocity or force exerted by the fingers.

Nowadays it is possible to build mechanically advanced prostheses, for example the *Otto Bock SensorHand*, *Otto Bock Michelangelo Hand* or the *Touch Bionics i-LIMB*. Dexterous robotics hand include the DLR prosthetic hand, the DLR Hit Hand (Liu, et al., 2008) and the *Shadow Robot Hand*.

Ultrasound imaging, to control a one-degree-of-freedom hand prosthesis, has been described in the work by Zheng, et al. (2006), Chen, et al. (2010) and Guo, et al. (2011). Here the focus lies on determining the muscle contraction of extensor muscles and is therein named sonomyography. A high correlation between the wrist extension angle and the change in size of the projection of the muscle itself is reported. As far as known, this method has not yet been applied to all fingers or the whole hand system.

As opposed to that, the method presented in this thesis is more concentrating on image processing and machine learning. Another difference is, that this method does not need a detailed anatomical model. The only attempt so far that is modelling finger positions is done by Shi, et al. (2010). Significant differences among optical flow computations for finger flexion movements are reported, but not analysed in detail.

3 Problem and Solution Approach

In this section, a short description of the problems in prosthetics is presented as well as the outline for the solution approach.

A state of the art approach for controlling prosthetic hands is using surface EMG, a technique by which muscle activation potentials are gathered by electrodes placed on the patient's skin. These potentials can be used to track what muscles the patient is activating, and with what force. Still, EMG signals suffer from a number of problems, for example optimal placement of the electrodes, signal drifting and signal change due to sweat formation, muscular fatigue and cross-talking among deep and superficial muscles. The dexterity of active hand prosthetics is limited not only due to the limited availability of dexterous prosthetic hands, but mainly due to limitations in interfaces (Castellini and van der Smagt, 2009).

In order for the system to be applicable on subjects, certain qualifications have to be fulfilled. For example it has to be ensured that enough muscular potential activity exists in the subject's forearm. Furthermore, the system has to work on different subjects with variations in musculoskeletal structures. Other points that have to be considered is the low signal-to-noise ratio of medical ultrasonic images and the inevitable movement of the ultrasound transducer.

The hereby presented approach tries to overcome these challenges. The process operation is as follows:

- Gathering ultrasound images of the forearm
- Gathering finger position via data glove
- Image processing and feature extraction
- Data synchronisation and normalisation
- Linear regression analysis
- Prediction of finger positions

In all the experiments described in this thesis, six degrees of motion are considered, namely flexion and extension of the four fingers (pinkie, ring, middle, and index) as well as thumb adduction and abduction. Additionally, thumb flexion and extension, which is roughly equivalent to thumb rotation. The choice of these six motions is also motivated by the fact that they are enforced by most advanced hand prostheses.

4 Theoretical Foundation

In this chapter the basics for ultrasonography, human anatomy, machine learning and image processing are presented.

4.1 Ultrasonography

Ultrasound imaging as a diagnostic device is widely used to investigate the interior of the human body. This section shall introduce the technical and imaging principles of medical ultrasonography.

4.1.1 Requirements

There are three requirements for diagnostic medical ultrasound: generation of a beam, reception of the returning echo, and processing of the signal for display. All types of ultrasonic scanning equipment have these basic features. Changes in the reception and in the processing, analysis, and display of returning echo signals differentiate one scanning mode from another (Hedrick, 1995).

4.1.2 Technical Principle

Basic Ultrasound Physics

Sound is mechanical energy that is transmitted through a medium. Periodic changes in the pressure of the medium are created by forces acting on the molecules, causing them to oscillate about their mean or average position. Sound waves are those pressure changes that the human ear can detect and oscillate at frequencies of 20 to 20,000 Hz. Ultrasound is defined as high-frequency mechanical waves with frequencies of greater than 20,000 Hz which humans cannot hear.

The periodic changes in pressure when vibrating molecules interact with neighbouring molecules are conveyed from one location to another. Sound waves require an elastic deformable medium for propagation which can be gas, liquid or solid. There is no flow of particles, the effect of compression and rarefaction is transmitted over long distances because of neighbour-to-neighbour interactions.

Elasticity refers to the ability of an object to return to its original shape and volume after a force is no longer acting on it. The amount of the change in volume and shape depends on the force and the elastic properties of the object. An ultrasound wave travelling through soft tissue causes elastic deformations by the separation and and compression of neighbouring molecules.

The velocity of sound through a medium is inversely proportional to the square root of the density of the medium. Thus ultrasound has a greater velocity in air (low density), than in bone (high density). The velocity of sound is also inversely proportional to the compressibility of the medium. The low compressibility of dense material (e.g. bones) predicts the high velocity of sound in bone. By contrast, air is easily reduced in volume and therefore the velocity of sound in air is low. Although exceptions can be cited, for materials of interest (air, fat, soft tissue, plastic, bone) the statement that, as density increases, the velocity

of sound increases, is generally true. Table 4.1 shows properties of different media (Hedrick, 1995).

Material	Density [kg/m ³]	Velocity [m/s]	Acoustic Impedance [kg/(m ² s)]
Air	1.2	330	0.0004
Water (20°)	1000	1480	1.48
Average Tissue	1060	1540	1.63
Muscle	1080	1580	1.70
Fat	952	1459	1.38
Blood	1057	1575	1.62
Bone	1912	4080	7.8

Table 4.1: Properties of different media (Hedrick, 1995)

In general, waves are divided into two basic types: longitudinal and transverse. Sound waves are longitudinal and those in which particle motion is along the direction of the wave energy propagation, whereas transverse waves are those in which the motion is perpendicular. Bone is the only biological tissue that can cause the production of transverse waves, which are sometimes referred to as shear waves or stress waves.

The velocity of ultrasound remains constant for a particular medium. The velocity (c) is equal to the frequency (f) times the wavelength (λ) (formula 4.1):

$$c = f\lambda \quad (4.1)$$

Because the velocity is constant, increasing the frequency causes the wavelength to decrease. The wavelength of ultrasound affects the axial resolution of an ultrasound system (Hedrick, 1995).

Interactions of Ultrasound with Tissue

For diagnostic ultrasound the recorded image is typically based on reflected energy. The single device that generates the ultrasound wave and, subsequently, detects the reflected energy is the transducer. An ultrasound wave is directed into the body to interact with tissues in accordance with the characteristics of the targeted tissues. The results of these interactions are recorded for diagnosis in the form of reflected ultrasound waves. The types of interactions that occur are similar to the wave behaviour observed with light: reflection, refraction, scattering, diffraction, divergence, interference, and absorption. With the exception of interference, all these interactions reduce the intensity of the beam, termed attenuation (Hedrick, 1995).

Reflection

The major interaction of interest for diagnostic ultrasound is reflection. If a sound beam is directed at right angles to a smooth interface larger than the width of the beam, it will be partially reflected toward the sound source (see figure 4.1). These interfaces (boundaries between different tissue types) are called specular reflectors.

The angle of reflection of a sound beam is equal to the angle of incidence. To obtain maximum detection of the reflected signal, the transducer is orientated so the generated sound beam will

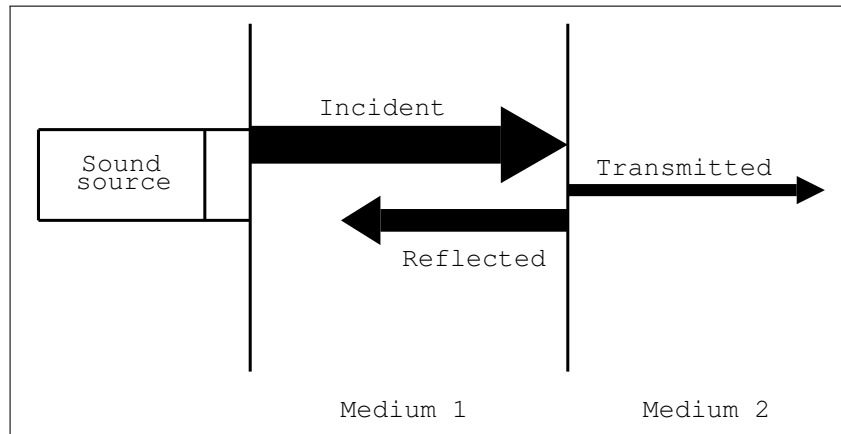


Figure 4.1: Partial reflection of an ultrasound beam (adopted from Hedrick, 1995)

strike the interface perpendicularly. Whereas in classical mechanics momentum is equal to mass times velocity, in ultrasound density replaces the mass. The velocity is that of sound in the medium. The product of density (ρ) times velocity (c) is called the acoustic impedance (Z)(formula 4.2):

$$Z = \rho c \quad (4.2)$$

The quantity is a measure of the resistance to sound passing through the medium and similar to electrical resistance. Acoustic impedance is expressed as kilograms per square meter per second ($kg/(m^2s)$). Table 4.1 lists the acoustic impedances for several materials of interest. As one can see, high-density materials give rise to high-velocity sound waves and therefore high acoustic impedances. If the acoustic impedance is the same in one medium as in another, sound will be readily transmitted from one to the other. It is primarily the change in acoustic impedance at a biological interface that allows visualisation of soft tissue structures with an ultrasonic beam. The amount of reflection is a function of the surface only. Therefore sound is reflected at the interface, regardless of the thickness of the material from which it is reflected. The reflection coefficient can be expressed as in formula 4.3:

$$R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2 \quad (4.3)$$

As one can see, it does not matter which impedance is the larger or smaller for two materials composing the interface. If the acoustic impedance difference is small, the magnitude of the reflecting wave will be small. If the acoustic impedance difference is large, such as in bone compared to soft tissue, a large fraction of sound will be reflected. Only little of the transmitted beam will penetrate structures behind the bone, and much will return to the detector (Hedrick, 1995).

Table 4.2 lists the percentage reflections at interfaces of varying composition.

Even if the air layer between a transducer and the skin is extremely thin, nearly total reflection occurs at the air-skin interface. During scanning, coupling gel is used to eliminate this. The gel also serves to reduce friction between the transducer and the skin.

When a device that produces and detects ultrasonic waves scans a patient, multiple interfaces are encountered along the path of travel of the ultrasound beam. A percentage of the incident

Interface	Reflection [%]
Soft tissue - air	99.9
Soft tissue - bone	43.0
Soft tissue - fat	0.69
Soft tissue - muscle	0.04

Table 4.2: Percentage reflection at different interfaces (Hedrick, 1995)

beam is reflected and transmitted at each interface. A series of echoes is subsequently detected. If losses by reflection only are considered, the relative intensities of these echoes depend on the acoustic impedance mismatch at the interface that originally created the echo and on the transmission of sound energy through the interfaces along the path of travel to and from the transducer. On the return path each interface allows a fraction of the echo energy to pass through toward the transducer (see figure 4.2).

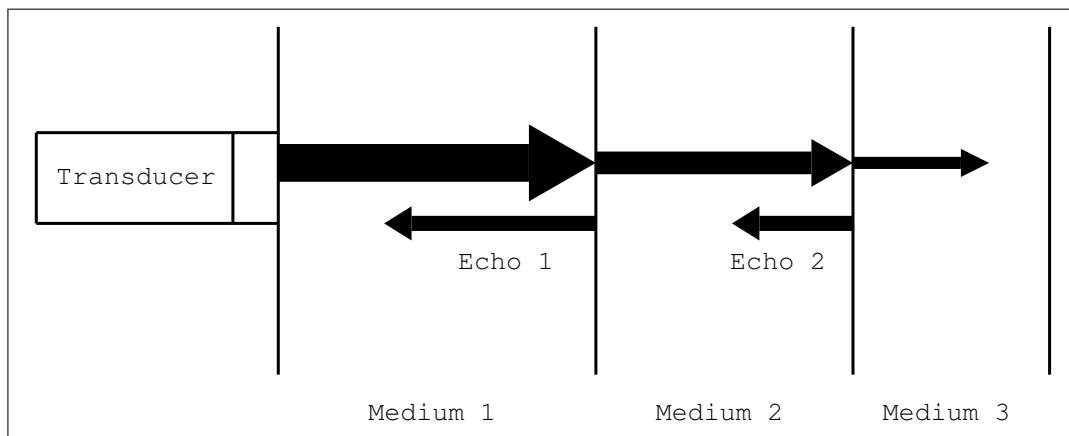


Figure 4.2: Series of echoes of an ultrasound beam (adopted from Hedrick, 1995)

A large smooth surface acts as a mirror to form a well-defined redirected beam. A large rough-surfaced interface deflects the ultrasound beam in multiple directions. Since the interface is not flat, the sound beam strikes the interface with various angles of incidence, which gives rise to differing angles of reflection. This is called diffuse reflection and weakens the echo returning to the transducer (Hedrick, 1995).

Scattering

Another important interaction between ultrasound and tissue is scattering, or non-specular reflection. It occurs because the interfaces are small, less than several wavelengths across. Each interface acts as a new separate sound source, and sound is reflected in all directions. These non-specular reflections have a strong frequency dependence, which may make them useful in characterising tissue. When the frequency is changed the altered scattering of the sound beam may provide important information for differentiating tissue types (Hedrick, 1995).

Reflectivity

Many factors influence the fraction of incident intensity that is reflected at an interface toward the transducer: the acoustic impedance mismatch, the angle of incidence, the size of the structure compared with the wavelength, the shape of the structure, and the texture of the surface of the interface. The combination of these factors is described by the term reflectivity.

Differences in reflectivity are partially responsible for the patient-to-patient variations that can be observed when performing a particular type of examination. Ultrasound imaging systems are capable of detecting extremely small changes in reflectivity on the order of one in a million (Hedrick, 1995).

Refraction

Another interaction that occurs between ultrasound and tissue is refraction. If the ultrasound beam strikes an interface between two media at an angle of 90 degrees, a percentage will be reflected back to the first medium and the rest will be transmitted into the second medium without a change in direction. If the beam strikes the interface at an angle other than 90 degrees, however, the transmitted part will be refracted or bent away from the straight-line path. Refraction of sound waves obeys Snell's law, which relates the angle of transmission to the relative velocities of sound in the two media. Although refraction is not a major problem in diagnostic ultrasound, under certain conditions the bending of the sound beam can cause artefacts in diagnostic images (Hedrick, 1995).

Interference

Sound waves demonstrate interference phenomena or the superposition of waves. If waves with the same frequency are in phase, they undergo constructive interference which results in an increased amplitude. If they are out of phase, they undergo destructive interference. Every combination, from completely constructive to completely destructive interference, can occur, resulting in a complex wave summation. This interference is important in the design of an ultrasonic transducer because it affects the uniformity of the beam intensity throughout the ultrasonic field. Focusing of the ultrasound beam in real time imaging is based on the principle of wave interference (Hedrick, 1995).

Absorption and Attenuation

Absorption is the only process whereby sound energy is dissipated in a medium. All other modes of interactions (reflections, refraction, scattering, and divergence) decrease the ultrasonic beam intensity by redirecting the energy of the beam. Absorption is the process whereby energy is transformed into other energy forms, primarily heat. The absorption of an ultrasonic beam is related to the beam's frequency and to the viscosity and the properties of the medium.

Attenuation includes the effects of both scattering and absorption in the characterisation of amplitude reduction as the ultrasound wave propagates through a medium. Attenuation is described by an exponential function (formula 4.4):

$$A = A_0 \exp(-az) \quad (4.4)$$

A is the peak amplitude of the beam at distance z and A_0 the original peak amplitude of the beam. The attenuation coefficient (a) is given by the sum of the scattering coefficient (a_s) and

the absorption coefficient (α) (formula 4.5):

$$a = a_s + \alpha \quad (4.5)$$

The various coefficients quantitate the respective fractional losses in amplitude per unit length from absorption, scattering, and both processes together (Hedrick, 1995).

Intensity

The intensity of an ultrasonic beam is the physical parameter that describes the amount of energy flowing through a cross-sectional area per second. It is the rate at which the energy is transmitted by the wave over a small area. For ultrasound, increasing intensity means, the distribution of particles within the compression regions becomes more dense, acoustic pressure is higher, length of particle oscillation increases and the maximum particle velocity is greater. The intensity of the ultrasound beam decreases as the beam propagates tissue. The transmitted intensity and the rate of intensity loss influence the ability of a scanner to observe weakly reflecting structures. Frequency, wavelength, and velocity of an ultrasonic beam are not affected by a change of intensity. The study of potential biological effects is linked to intensity. Since the particle velocity and length of displacement are dictated by intensity, a high-intensity ultrasound wave is more disruptive than a low-intensity ultrasound wave is to living systems.

Although no standard reference intensity for ultrasound has been established, a useful method for determining the reduced intensity of a beam is to make relative measurements that compare the value at one point with a reference intensity at another point. Relative measurements are usually made and given in decibels (dB) (Hedrick, 1995).

Penetration

High-frequency sound waves are attenuated faster than low-frequency sound waves. Thus the ability to penetrate tissue is reduced at higher frequencies. In addition, a reflector positioned at progressively greater depths generates progressively lower-intensity returning echoes. Table 4.3 shows some intensity attenuation factors for human tissues at a frequency of 1 MHz.

Material	dB/cm
Blood	0.18
Fat	0.6
Muscle (across fibers)	3.3
Muscle (along fibers)	1.2
Water	0.0022

Table 4.3: Attenuation of human tissues at 1MHz (Hedrick, 1995)

Determination of Echo Intensity

In ultrasound imaging the transducer sends out a pulsed wave and subsequently detects the returning echo. Intensity loss occurs by attenuation of the transmitted wave going out to the interface and also by attenuation of the reflected wave coming back toward the transducer from the interface. The rate of attenuation depends on tissue type and wave frequency. The intensity

of the detected echo from an interface is diminished by increasing the distance of travel. The large reduction attributed to attenuation in bone is the major reason why bone generally avoided during an ultrasound examination. Additionally there is a large amount of reflection at any soft tissue-bone interface (Hedrick, 1995).

Echo Ranging

In diagnostic ultrasound, reflections of the sound beam from interfaces along the ultrasonic path are of primary interest. An ultrasound wave is transmitted into the body, strikes an interface, and is partially reflected to the transducer, as determined by the percentage reflection formula (see formula 4.3). The reflected waves arising from the various acoustic impedance mismatches result in ultrasonic detection of interfaces within the body. The distance (z) travelled by the sound beam is specified by the formula 4.6, where c is the velocity and t the time of travel:

$$z = ct \quad (4.6)$$

If the velocity of ultrasound and the elapsed time from the original send pulse to the detection of the return echo are known, the distance to an interface can be determined with formula 4.6.

A system that can generate an ultrasonic pulsed wave and detect the reflected echo after a measured time permits the distance to an interface to be determined. This technique is called echo-ranging, a concept that formed the basis of sonar (sound navigation and ranging) developed during World War II.

For echo ranging to delineate the position of an interface correctly, certain conditions must hold: First, the ultrasound wave must travel directly to the interface and back to the transducer along a straight-line path and second, the velocity of sound must remain constant along the path of travel. The principle of echo ranging can be seen in figure 4.3 (Hedrick, 1995).

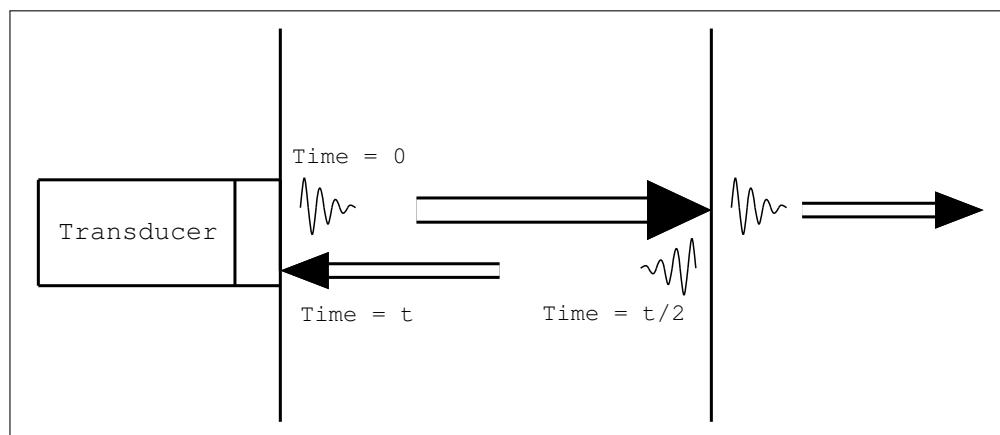


Figure 4.3: Principle of echo ranging (adopted from Hedrick, 1995)

Basic Scanner

Although the basic scanner is electronically complex, a block diagram in figure 4.4 describes the main system components and operation of the unit. All the subsystems are controlled by a master synchroniser. An electronic signal from the master synchroniser to the transmitter initiates the

process. A pulsed ultrasound beam is generated and sent into the body. It undergoes various interactions and when an interface is encountered, a fraction of the beam's energy is reflected back toward the transducer. The transducer, acting as a receiver rather than a transmitter, converts the ultrasound wave to an electronic signal that is processed and displayed. To correlate the amplitude of the induced signal with the depth of the interface, the time interval between the time that an ultrasound pulse leaves the transducer and the time the echo returns must be known. All the subsystems must be activated simultaneously by the master synchroniser to ensure proper timing. For each system the clock in the master synchroniser measures the elapsed time from generation of the ultrasound pulse to reception of the echoes (Hedrick, 1995).

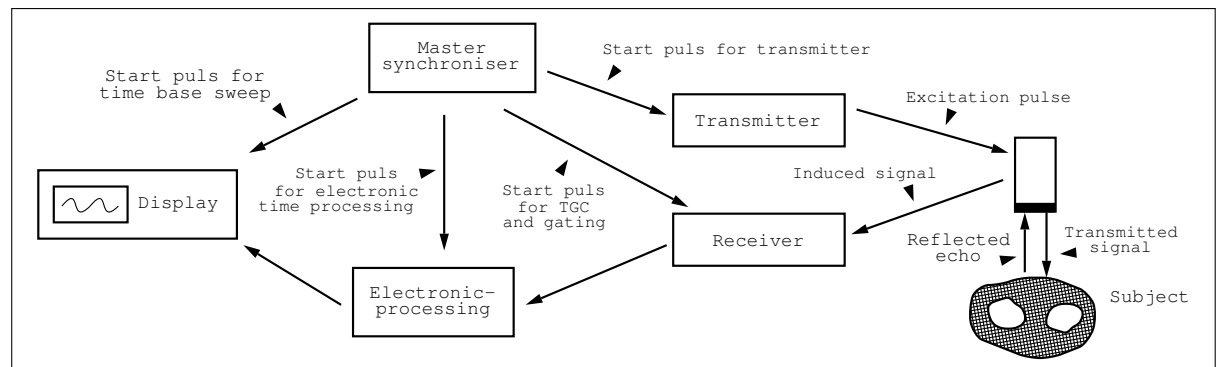


Figure 4.4: Main system components of a basic scanner (adopted from Hedrick, 1995)

4.1.3 Static Imaging Principles

Sonar played a major role in the development of medical diagnostic ultrasound instrumentation, particularly in A-mode scanning (Hedrick, 1995).

A-mode Scanning

A-mode (amplitude-mode) scanning is based on the echo-ranging principle, similar to sonar. A pulsed ultrasound wave is directed into the patient's body, and the echoes generated at various interfaces are detected. Only structures that lie along the direction of propagation are interrogated. This sampling according to the beam path is called the line of sight or scan line. A-mode displays the amplitude of the signal as a spike in the vertical dimension versus depth or time of the signal in the horizontal dimension. Time gain compensation attempts to correct for attenuation loss. Multiple interfaces encountered along the sampling direction are detected by a series of echoes (Hedrick, 1995).

B-mode Scanning

In B-mode imaging the amplitude of the signal is represented by the brightness of a dot. The A-mode spike is converted into a dot and the amplitude or strength of the signal is designated by the dot's brightness. The position of the dot represents the depth (time) of the interface from the transducer. Instead of acquiring data from a single scan line or line of sight, the aim is to construct a two-dimensional image of the area of interest. This is accomplished by

compound B-scanning, whereby multiple sets of dots are combined to delineate the echo pattern from internal structures within the body. The superimposition of multiple scan lines creates a composite two-dimensional image, which has the advantage of portraying the general contour of the patient and the internal structure. Compound B-mode scanning produces a static image that can be envisioned as a stop-action photograph of the reflecting surfaces.

Two problems with static B-mode scanning not present with A-mode are registration and storage of the scan-line information (Hedrick, 1995).

Two applications of B-Mode ultrasound imaging are displayed in figure 4.5. Prenatal screening of a foetus on the left side and visualisation of musculoskeletal structures on the right side.

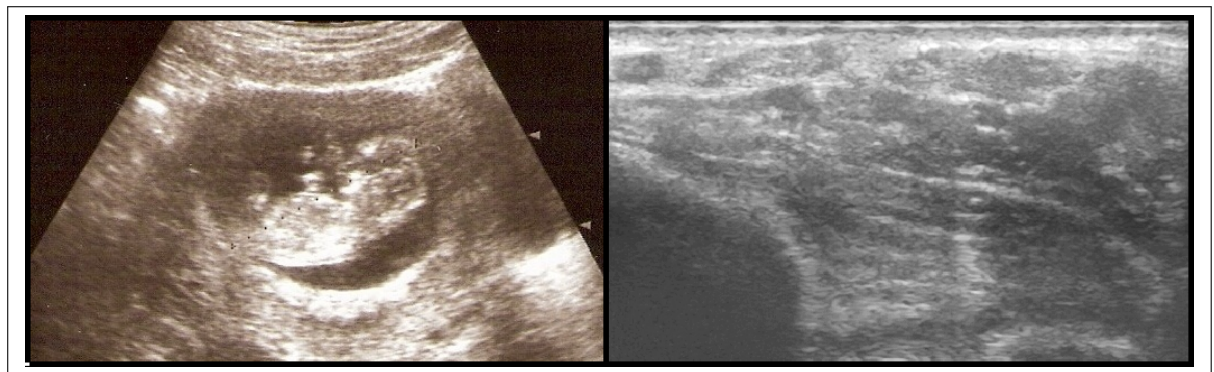


Figure 4.5: Two applications of B-mode scanning: Prenatal screening of a foetus on the left and the visualisation of musculoskeletal structures (human forearm) on the right

4.1.4 Real-Time Ultrasound Imaging Principles

Two-dimensional ultrasound imaging is performed with either static B-mode grayscale units or real-time grayscale scanners. In real-time scanning the displayed image is continuously and rapidly updated with new scan data as the beam is swept repeatedly throughout the field of view. The rate at which new information is displayed can be 30 or more frames per second. The rapid frame rates in real-time imaging eliminate difficulty with motion artefacts, which greatly compromise static imaging (Hedrick, 1995).

Principle

In real-time grayscale imaging the ultrasound beam is first directed along on line of sight, and after the echoes are received it is automatically moved to a new line. A single image is formed by sweeping through the entire region. This process is repeated to produce successive images of the region. As motion becomes more rapid within the field of view, a faster frame rate is necessary to display the structures. The time for the pulse to travel to the depth of interest and back to the transducer imposes a restriction on the frame rate. Frame rates of 5 to 40 images per second are available. The number of scan lines in a frame is usually between 50 and 200 depending on scan conditions. Commonly 120 to 150 scan lines are used (Hedrick, 1995).

Frame Rate Limitations

The maximum frame rate (FR) in frames per second (fps) is given by the following equation (formula 4.7):

$$FR = \frac{c}{2RN} = \frac{PRF}{N} \quad (4.7)$$

where c is the velocity of ultrasound in the medium, R the depth of interest, N the number of lines of sight per frame (lpf), and PRF the pulse repetition frequency. This equation indicates that, if the scanning depth or number of lines of sight are increased, the maximum frame rate must decrease. The number of frames per second is ultimately limited by the velocity of ultrasound in tissue (1540 m/s). The decreased penetration enables a higher frequency transducer to be used for better axial resolution (Hedrick, 1995).

Linear Arrays

An important advance in medical diagnostic ultrasound instrumentation occurred in the late 1970s with the development of electronic real-time instrumentation. These systems produce high-quality images with high frame rates and many lines of sight. The transducer of such systems consists of multiple crystals arranged in a straight row. Each crystal produces an ultrasound beam and then receives the returning echoes for data collection along one line of sight. The crystals are activated in a sequential fashion to form the individual lines of sight. The number of crystals in the array determines the number of lines of sight of each image. The first crystal is fired, and then a time delay for collection of the returning echoes are imposed before the next crystal is fired. The duration of the delay is set by the maximum scanning depth. The rate at which this sequence of firing all crystals is repeated determines the frame rate. In segmental linear arrays several crystals can be grouped and fired at the same time. Figure 4.6 illustrates the timing pattern for a sequential linear array. The duration from time t_0 to time t_1 depends on the depth. N is the number of lines per frame (Hedrick, 1995).

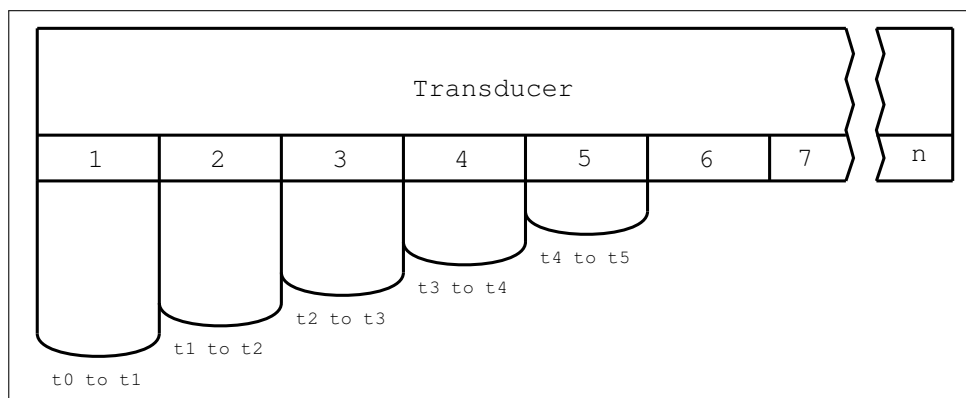


Figure 4.6: Timing pattern for a sequential linear array (adopted from Hedrick, 1995)

Principle of Focusing

Ultrasound beams are focused to improve the lateral resolution and sensitivity. Electronic focusing involves the superimposition of ultrasound waves. Each crystal produces a particular wave pattern,

and the overall pattern derived from a group of crystals is the summation of all the wave patterns from the individual crystals. Electronic focusing is accomplished by offsetting the firing of various crystals in a group by a small delay. This delay is small compared with the time required for the sound beam to travel to the depth of interest. For the waves from each crystal to be in phase at the point of interest, constructive interference must occur. Each wavefront must arrive at the point of interest at exactly the same time. The ultrasound beam thus is reinforced or focused in the length direction. The firing sequence is repeated for the next image in the real-time acquisition. Improved spatial detail is thus obtained within the focal zone (Hedrick, 1995).

4.1.5 Risks

Diagnostic ultrasound imaging is considered to be risk-free and no harms whatsoever are reported. For example, in the World Health Organisation Technical Report series (WHO Study Group, 1998), it says:

"Ultrasonography has established an enviable safety record. There has not yet been any published report of harmful biological effects due to diagnostic ultrasound, on either patients or operators".

4.2 Anatomy of the Human Arm

In order to understand how certain muscle activations are related to the corresponding finger movements, an anatomical outline of the human forearm is given in this section. The general principle of operations as well as names and positions of the relevant muscles are explained.

The human muscular system is made up on the voluntary skeletal muscle organs of the body. Each of these organs is a complex of skeletal muscle tissue, a connective tissue framework, and a rich supply of blood vessels, lymphatics, and nerves. They produce all the voluntary movements, hence are one of the most important systems of the body. They are essential to the performance of a great variety of body functions (Gray, 2008).

Locomotion is one of the most obvious functions of this system. It enables humans to creep, walk, and run. But beyond all of this, among the most important and unique tools are the hands - miracles of quick and coordinated actions - capable of performing a vast assortment of activities, from swinging a club to fingering the keys of a piano. Among the great events of man's evolution was the attainment of the bipedal condition by which the hands were freed from locomotion for other uses (Gray, 2008).

4.2.1 Terms of Direction and Terms of Movements

In order to communicate effectively about the human body, anatomists have agreed upon the rule of referring all descriptions to the so called anatomical position. Table 4.4 lists some relevant terms of direction.

The terms proximal and distal are used to denote distance from source or attachment of a structure or from the centre of the body. For example, the proximal end of a finger is at its

Term	Direction
Distal	Further from the body
Proximal	Nearer to the body
Posterior, dorsal	Back
Anterior, palmar, volar	Front or belly side
Medial	Nearer the midline
Lateral	Away from the midline

Table 4.4: Terms of direction (Crouch, 1978)

point of attachment to the hand and its free end is distal. The terms superficial and deep refer to relative distances from the body surface. The anterior surfaces of the hands or forearms are called volar or palmar. Their opposite surfaces are called posterior (Crouch, 1978).

The basic active movements such as flexion-extension, abduction, adduction, rotation, and circumduction are listed in table 4.5.

Term	Movement
Flexion	Decrease in angle between two bones
Extension	Increase in angle between two bones
Adduction	Movement toward the central axis of the body part
Abduction	Movement away from the central axis of the body part
Rotation	Circular movement
Circumduction	Circular movement
Supination	Lateral rotation
Pronation	Medial rotation

Table 4.5: Terms of movements (Crouch, 1978)

Pronation is the medial rotation of the forearm which brings the palm of the hand into the backward position. The radius, in this case, lies diagonally across the ulna. Supination is the lateral rotation of the forearm which brings the palm of the hand into the forward position. In this position the radius and the ulna are parallel (Crouch, 1978).

The hand also has a well-recognised position of rest, with the wrist in extension and the digits in some degree of flexion. The precise position of the thumb in the position of rest appears to be rather variable. Typically it is considered as the midpoint between maximal palmar abduction and maximal retroposition (Gray, 2008).

This position maybe not be confused with the resting position (also called reference position) described in the experiments in section 6. Whereas in the position of rest all the muscles are relaxed, in the reference position the fingers and the thumb are almost fully extended. The reference position can be seen in figure 6.1, movement number 1.

4.2.2 Muscles of the Human forearm and Hand

The muscles of the forearm are many and varied. Fortunately, their names often include words descriptive of their chief functions and sometimes indicate their origins of insertion. These muscles arise mostly from the distal end of the humerus, a few from the radius and ulna. Their bellies form the fullness of the proximal end of the forearm. Distally they become tendinous and many pass into the hand (Crouch, 1978).

Muscles can only pull in one direction. For this reason, most muscles of the human body are arranged in pairs. When one muscle contracts, to flex a finger for example, its counterpart remains relaxed. An opposable thumb requires a different system of control from the other digits. Since the metacarpal is much more mobile than in the digits, muscles are needed to control the extra degree of freedom of movement (Crouch, 1978).

Normally functioning body movements, such as the flexion of the knee or of the elbow, do not involve just individual muscles but rather groups of muscles. In one kind of movement of a part, a given muscle functions in a certain way; in another kind of movement of the same part the same muscle plays a different role. If a muscle is the primary agent in the production of a desired movement, it is called a prime mover. If the same muscle serves to counteract or slow the action of another prime mover, it is called an antagonist. If it serves to steady a movement or to eliminate an undesired movement of a joint, it is said to be a synergist. Muscles in general are arranged in opposing positions on the various parts of the body, flexors-extensors, abductors-adductors. Further, when the fist is clenched, the flexors of the fingers and thumb contract, and if this action were unopposed, the wrist would bend too, and interfere with the fist-clenching. Flexion of the wrist in this situation is opposed by the extensors of the wrist which aid the fist-clenching operation. In this case the extensors are acting as synergists (Crouch, 1978).

Table 4.6 lists the most important muscles of the human forearm. These muscles are mainly needed to move the fingers and the wrist. In all the experiments described in this thesis, the flexor muscles are of the utmost significance (see experiment description in section 6).

The capacity and versatility of the human hand is due largely to the opposable thumb, and to the numerous joints of the wrist and even those of the forearm, arm, and shoulder which enable one to use the hand in many positions. It follows, of course, that the hand must have many muscles (also called interior muscles) within itself to supplement those that enter it from neighbouring parts or indirectly influence its action (Crouch, 1978).

The interior muscles of the hand are listed in table 4.7 for the sake of completeness. They play no role in the experiments because they are not observed by the ultrasonic system but the high number of muscles give an insight into the complexity and intricacy of finger movements.

The intrinsic muscles of the hand are organised into three groups plus a superficial muscle. Flexor pollicis brevis, abductor pollicis brevis, opponens pollicis and adductor pollicis all act on the thumb and are known collectively as the thenar muscles. Abductor digiti minimi, flexor digiti minimi brevis and opponens digiti minimi all act on the pinkie and are known collectively as the hypothenar muscles. Interossei and lumbricals act on the fingers. Palmaris brevis is a superficial muscle that lies beneath ulnar palmar skin (Gray, 2008).

Muscle	Action
Flexor carpi radialis	Flexes and abducts wrist
Flexor carpi ulnaris	Flexes and adducts wrist
Flexor digitorum superficialis	Flexes middle phalanges of each finger
Flexor digitorum profundus	Flexes distal phalanges of each finger
Flexor pollicis longus	Flexes distal phalanx of thumb
Flexor digiti minimi brevis	Flexes pinkie
Palmaris longus	Flexes wrist
Brachioradialis	Flexes forearm
Extensor carpi radialis longus	Extends and abducts wrist
Extensor carpi radialis brevis	Extends and abducts wrist
Extensor digitorum	Extends phalanges
Extensor indicis	Extends index finger
Extensor carpi ulnaris	Extends and adducts wrist
Extensor pollicis brevis	Extends thumb
Extensor pollicis longus	Abducts thumb
Extensor digiti minimi	Extends of joints of pinkie
Pronator quadratis	Pronates forearm

Table 4.6: Muscles of the forearm that move the wrist and fingers (Montgomery, 1981)

Muscle	Action
Abductor pollicis brevis	Abducts thumb
Flexor pollicis brevis	Flexes proximal phalanx of thumb
Adductor pollicis	Adducts thumb
Opponens pollicis	Opposes thumb
Flexor digiti minimi brevis	Flexes pinkie
Abductor digiti minimi brevis	Flexes pinkie
Opponens digiti minimi	Opposes pinkie
Lumbricals	Flexion of fingers
Palmar interossei	Adduct fingers towards middle finger
Dorsal interossei	Abduct fingers away from middle finger
Palmaris brevis	May contribute to palmar grip

Table 4.7: Interior muscles of the hand that move the fingers (Montgomery, 1981)

In order to see where exactly in the forearm the muscles are located, two cross sections of the right forearm are displayed in figure 4.8 and figure 4.10. Figure 4.7 indicates the exact position of the cross sections. These positions are approximately the same positions of the ultrasonic transducer in the experiments.

In figure 4.8 (cross section A, at the distal end of the right forearm) one can see the muscles and tendons of the forearm. Muscles are represented by gray, hatched areas whereas tendons are represented by plain, white areas. One can see that the flexor muscles and tendons lie on

the palmar side of the arm and the extensor muscles on the dorsal side. The two groups are separated by the ulna and radius and the pronator quadratis muscle. Nerves, arteries and veins (white, dashed areas) play no role in the examinations.

The ulna is the longer of the two bones of the forearm. It lies medial and parallel to the radius, the lateral bone of the forearm (Crouch, 1978). The two bones are connected in life by an interosseous membrane, a broad, thin, collagenous sheet that provides attachments for the deep forearm muscles. (Gray, 2008) .

The most relevant muscles for the experiments are the flexor muscles, especially the flexor digitorum superficialis and flexor digitorum profundus. In figure 4.9 the ultrasound image is compared to the corresponding part of the anatomical image. This figure also illustrates what part is examined by the ultrasonic system. In the ultrasound image, the pronator quadratis muscle (dark area circumscribed by the two inclined, bright lines on the lower right side) is clearly visible. This muscle is not activated when the fingers are moved and usually does not change during an experiment. In the ultrasound image one can also guess the structures of the other flexor muscles.

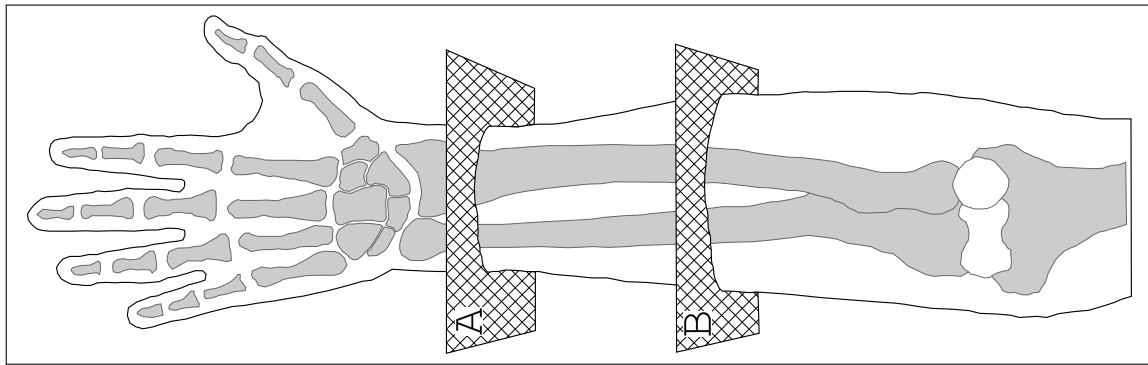


Figure 4.7: Position of cross section A and cross section B; right forearm (reproduced from Netter, 2006)

Figure 4.10 illustrates another cross section (cross section B, mid-arm). One can see how the size and position of each muscle is changing. In comparison to cross section A the size of the muscles has increased and there are no more tendons. Also the pronator quadratis muscle is no longer visible because it is located at the distal end of the forearm.

The bright horizontal line in the middle of the ultrasound image on the right in figure 4.11 indicates the border between the flexor digitorum superficialis muscle and flexor digitorum profundus muscle. Also the border between the flexor muscles and the extensor muscles together with the interosseous membrane can be seen (inclined, bright line on the lower right side).

Besides muscles, tendons and the two bones ulna and radius, the human forearm consists of other soft tissues like fascia. The fascia enclose, or invest, various structures of the body. They vary in thickness and strength, and in the relative amounts of fat, collagenous and elastic fibers, and tissue fluid which they contain. They constitute a structural and functional system of the body (Crouch, 1978).

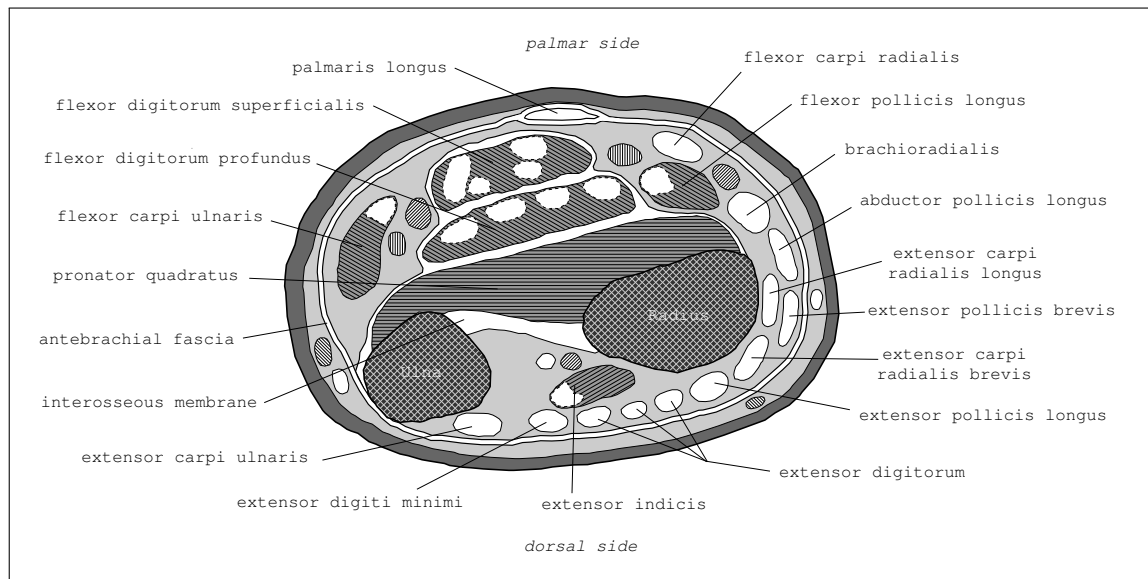


Figure 4.8: Cross section A: Anatomical structure of the right forearm; distal (reproduced from Netter, 2006)

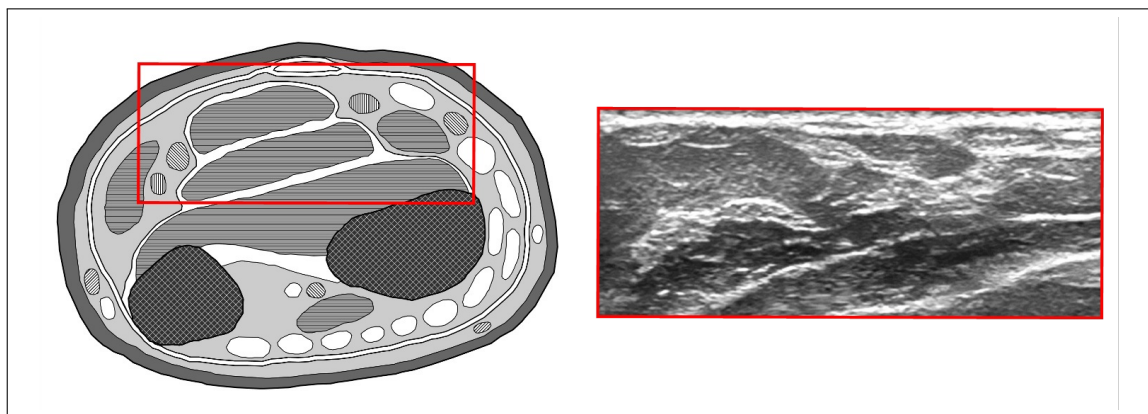


Figure 4.9: Anatomical image compared to ultrasound image (cross section A)

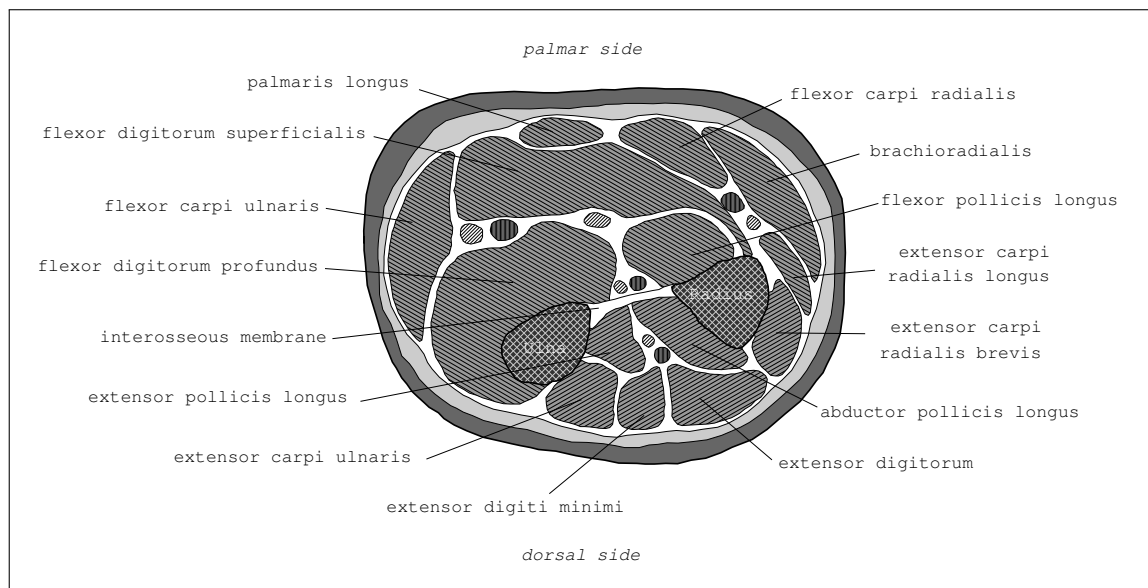


Figure 4.10: Cross section B: Anatomical structure of the right forearm (mid-arm)

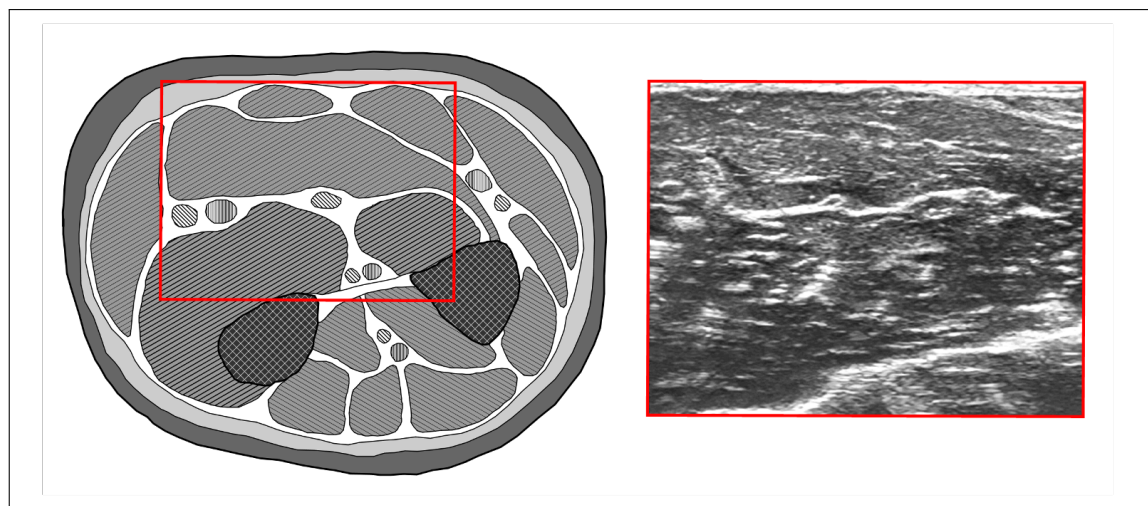


Figure 4.11: Anatomical image compared to ultrasound image (cross section B)

The antebrachial fascia is a continuation of the brachial fascia which forms a strong investment of the forearm. It is attached below to the distal parts of the radius and ulna and is continuous with the fascia of the hand. Through most of its length it is attached to the dorsal border of the ulna, thus closing off the palmar, or flexor, from the posterior, or extensor, compartments of the forearm. At the distal end of the forearm the fascia is thickened to form flexor and extensor retinacula, which serve to bind down the long tendons of insertion of the flexor and extensor muscles as they pass over this region (Crouch, 1978).

Muscle Variations

In this section, considerable muscle variations in the forearm that could have an effect in the experiments are discussed. Variations arise primarily from the human's genetic composition, an inheritance carried over from ancient origins. It is important to understand that no two living organisms are structurally or functionally identical - animals or plants (Bergman, 2006).

For instance, the palmaris longus is one of the most variable muscles in the body. It is absent about 10 percent of the population. It may consist solely of a tendinous band or the muscle may be double. Also partial or complete insertion into the fascia of the forearm, into the tendon of the flexor carpi ulnaris, and into the muscles of the little finger have been observed (Gray, 2008). In a study by (Sebastin, et al., 2005) no statistically significant difference was seen in the grip or pinch strength measurements between subjects who had a palmaris longus tendon and those who did not.

4.2.3 Bones of the Human Hand

The bones of the hand are displayed in figure 4.12. The skeleton of the hand is subdivided into three segments: the carpus or wrist bones, the metacarpus or bones of the palm, and the phalanges or bones of the digits (Gray, 2008).

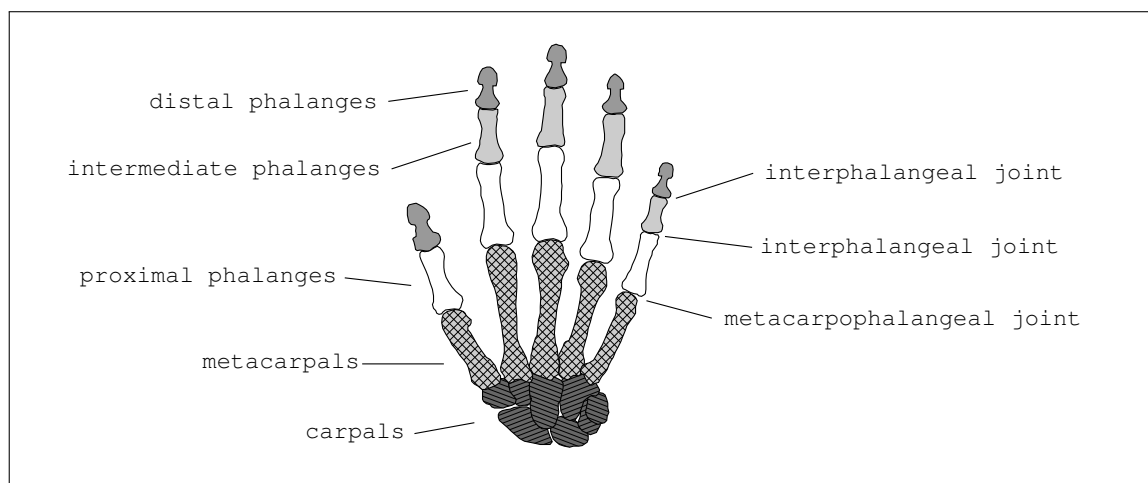


Figure 4.12: Bones of the human hand (adopted from Gray, 2008)

4.3 Machine Learning and Pattern Recognition

In order to be able to predict the finger positions correctly, the system that tries to do so has to be trained beforehand. This very crucial function is called machine learning. The fundamentals and different types of machine learning and pattern recognition shall be presented in this section.

The problem of searching for patterns in data is a fundamental one and has a long and successful history. For instance, the discovery of the empirical laws of planetary motion, or the discovery of regularities in atomic spectra. The field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such as classifying the data into different categories (Bishop, 2007).

4.3.1 Terminology

In the following section some of the terminology used in machine learning theory is introduced. Assumed that there is a function f and the task of the learner is to guess what it is, the hypothesis about the function to be learned is denoted by h . Both f and h are functions of a vector-valued input $\mathbf{X} = (x_1, x_2, \dots, x_i, \dots, x_n)$ which has n components. Furthermore h can be implemented by a device that has \mathbf{X} as input and $h(\mathbf{X})$ as output. Both f and h themselves may be vector-valued. The hypothesised function h is *a priori* selected from a class of functions \mathcal{H} . Sometimes f also belongs to this class or to a subset of this class (Nilsson, 1998).

Types of Learning

There are two major settings in which functions can be learned, unsupervised learning and supervised learning. In the latter one, the one that is used in all the experiments in this thesis, the values of f for the m samples in the training set Ξ are known. Assumed, a hypothesis h can be found that closely agrees with f for the members of Ξ , then this hypothesis will be a good guess for f , especially if Ξ is large (Nilsson, 1998).

Even though there is no correct answer in inductive learning, it is important to have methods to evaluate the result of learning. In supervised learning the induced function is usually evaluated on a separate set of inputs and function values for them called the testing set. A hypothesised function is said to generalise when it guesses well on the testing set. Both mean-squared-error and the total number of errors are common measures (Nilsson, 1998).

Cross Validation

Often the use of a test set is regarded as too wasteful of scarce classified data. If the training set is divided into two halves, it is possible to train with one half and test with the other. As the halves are independent samples, the resulting estimator is unbiased. Furthermore, by swapping the halves, again an unbiased estimator can be obtained. The average of the two estimators remains unbiased.

The drawback of this approach is that the estimate is an unbiased estimate of the performance using just half the data. At the expense of more computation, a better result can be achieved. If the Training set is randomly divided into V parts, then one piece can be used to test the

performance of the classifier trained on the remaining $(V-1)$ pieces. This is again unbiased and V can be averaged. For moderate V such as 5 or 10, the loss of performance from a smaller training set will usually be small enough, at the expense of V times the computation (Ripley, 2002).

Data Normalisation

Data normalisation is a useful step often adopted as a precaution when the feature values vary in different dynamic ranges. In the absence of normalisation, features with large values have a stronger influence on the cost function in designing the classifier. Data normalisation restricts the values of all features within predetermined ranges. A common technique is to normalise the features to zero mean and unit variance via linear methods. Assuming the training data contain N values of a specific feature x , the normalised values can be expressed as in formula 4.8

$$\hat{x}_i = \frac{x_i - \bar{x}}{\sigma}, \quad i = 1, 2, \dots, N \quad (4.8)$$

where \bar{x} and σ be the respective mean value and standard deviation (Theodoridis, et al., 2010).

Correlation Coefficient

The correlation coefficient (r or ρ) determines the correlation or linear dependence between two sets of variables. The value ranges from $\rho = 1$ for a perfect positive correlation to $\rho = -1$ for a perfect negative correlation. A complete absence of correlation is represented by $\rho = 0$.

The correlation coefficient for the two sets of variables x and y is denoted by

$$\rho_{x,y} = \frac{cov[x, y]}{\sqrt{var[x]} \sqrt{var[y]}} \quad (4.9)$$

where $cov(x)$ is the covariance of the variable x :

$$cov[x, y] = E[xy] - E[x]E[y] = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(x_y - \bar{y}) \quad (4.10)$$

with $E[x]$ as the expected value and $var[x]$ the variance of x , simply denoted by

$$var[x] = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (4.11)$$

(Bartsch, 2003).

Linear Dichotomies

Consider a set \mathcal{H} of functions, and a set, Ξ of patterns. One measure of the expressive power of a set of hypotheses, relative to Ξ is its ability to make arbitrary classifications of the patterns in Ξ . If there are m patterns in Ξ there are 2^m different ways to divide these patterns into two disjoint and exhaustive subsets. In other words there are 2^m different dichotomies of Ξ . If Ξ were to include all of the 2^n Boolean patterns, for example, there are 2^{2^n} ways to dichotomise them, and the set of all possible Boolean functions dichotomises them in all of these ways. But a subset \mathcal{H} of the Boolean functions might not be able to dichotomise an arbitrary set Ξ of m Boolean patterns in all 2^m ways. In general, if a subset \mathcal{H} of

functions can dichotomise a set Ξ of m patterns in all 2^m ways, then \mathcal{H} shatters Ξ (Nilsson, 1998).

The number of dichotomies achievable by a hyperplane depends on how the patterns are disposed. For the maximum number of linear dichotomies, the points must be in what is called general position. For $m > n$, a set of m points is in general position in an n -dimensional space if and only if no subset of $(n + 1)$ points lies on an $(n - 1)$ -dimensional hyperplane. When $m \leq n$, a set of m points is in general position if no $(m - 2)$ -dimensional hyperplane contains the set. Note that even though there are an infinite number of hyperplanes, there are, nevertheless, only a finite number of ways in which hyperplanes can dichotomise a finite number of patterns. Small movements of a hyperplane typically do not change the classifications of any patterns.

For each class \mathcal{H} there will be some maximum value of m for which \mathcal{H} shatters the m patterns. This maximum number is called the Vapnik-Chervonenkis dimension (VC-dimension) and is denoted by $\text{VCdim}(\mathcal{H})$ (Nilsson, 1998).

4.3.2 Linear Models for Regression

The goal of regression is to predict the value on one or more continuous *target* variables t given the value of D -dimensional vector \mathbf{x} of *input* variables. The simplest form of linear regression models are linear functions of the input variables. It is also possible to obtain a much more useful class of functions by taking linear combinations of a fixed set of non-linear functions of the input variables, known as *basis functions*. Such models are linear functions of the parameters, which gives them simple analytical properties, and yet can be non-linear with respect to the input variables.

Given a training data set comprising N observations together with corresponding target values, the goal is to predict the value of t for a new value of \mathbf{x} . In the simplest approach, this can be done by directly constructing an appropriate function $y(\mathbf{x})$ whose values for new inputs \mathbf{x} constitute the predictions for the corresponding values of t . More generally, from a probabilistic perspective, the aim is to model the predictive distribution because this expresses the uncertainty about the value of t for each value of \mathbf{x} . From this conditional distribution, predictions of t for any new value of \mathbf{x} can be made in such a way as to minimise the expected value of a suitably chosen loss function. A common choice of loss function for real-valued variables is the squared loss, for which the optimal solution is given by the conditional expectation of t .

Although linear models have significant limitations as practical techniques for pattern recognition, particularly for problems involving input spaces of high dimensionality, they have nice analytical properties and form the foundation for more sophisticated models (Bishop, 2007).

Least Mean Squares

In this chapter the least squares approach and its relation to maximum likelihood is discussed in more detail. First of all, to be able to understand the derivation of the formulas, one of the most important probability distributions for continuous variables, the *normal* or *Gaussian* distribution shall be introduced. For the case of a single real-valued variable x , the Gaussian distribution is defined by

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\} \quad (4.12)$$

which is governed by two parameters μ , called mean, and σ^2 , called the variance. The square root of the variance, given by σ , is called the standard deviation, and the reciprocal of the variance, written as $\beta = 1/\sigma^2$, is called the precision (Bishop, 2007).

A linear combination of input variables can be denoted by

$$y(\mathbf{x}, \mathbf{w}) = w_0 + w_1 x_1 + \dots + w_D x_D \quad (4.13)$$

where $\mathbf{x} = (x_1, \dots, x_D)^T$. The key property of this model is that it is a linear function of the parameters w_1, \dots, w_D (the parameter w_0 allows for any fixed offset in the data and is sometimes called a bias parameter). The model can also be denoted by

$$y(\mathbf{x}, \mathbf{w}) = w_0 + \sum_{j=1}^{M-1} w_j \phi_j(\mathbf{x}) \quad (4.14)$$

where $\phi_j(\mathbf{x})$ are known as basis function. By using non-linear basis functions, the function $y(\mathbf{x}, \mathbf{w})$ can be a non-linear function of the input vector \mathbf{x} . For the system described in this thesis, however, the vector $\phi(\mathbf{x})$ of basis functions is simply the identity $\phi(\mathbf{x}) = \mathbf{x}$. By denoting the maximum value of the index j by $M-1$, the total number of parameters in this model will be M . It is often convenient to define an additional dummy *basis function* $\phi_0(\mathbf{x}) = 1$ so that

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=0}^{M-1} w_j \phi_j(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) \quad (4.15)$$

where $\mathbf{w} = (w_0, \dots, w_{M-1})^T$ and $\phi = (\phi_0, \dots, \phi_{M-1})^T$ (Bishop, 2007).

Furthermore, the target variable t is given by a deterministic function $y(\mathbf{x}, \mathbf{w})$ with additive Gaussian noise so that

$$t = y(\mathbf{x}, \mathbf{w}) + \epsilon \quad (4.16)$$

where ϵ is a zero mean Gaussian random variable with precision (inverse variance) β . Thus it can be written as

$$p(t|\mathbf{x}, \mathbf{w}, \beta) = \mathcal{N}(t|y(\mathbf{x}, \mathbf{w}), \beta^{-1}) \quad (4.17)$$

Considered a data set of inputs $\mathbf{X} = \mathbf{x}_1, \dots, \mathbf{x}_N$ with corresponding target values t_1, \dots, t_N , the target variables t_n can be grouped into a column vector that is denoted by \mathbf{T} (not to be confused with a single observation of a multivariate target, which would be denoted by \mathbf{t}). Making the assumption that these data points are drawn independently from the distribution 4.17, the following expression for the likelihood function is obtained, which is a function of the adjustable parameters \mathbf{w} and β , in the form

$$p(\mathbf{T}|\mathbf{X}, \mathbf{w}, \beta) = \prod_{n=1}^N \mathcal{N}(t_n|\mathbf{w}^T \phi(\mathbf{x}_n), \beta^{-1}) \quad (4.18)$$

where 4.15 has been used. Taking the logarithm of the likelihood function, and making use of the standard form 4.12 for the univariate Gaussian leads to

$$\begin{aligned} \ln p(\mathbf{T}|\mathbf{w}, \beta) &= \sum_{n=1}^N \ln \mathcal{N}(t_n|\mathbf{w}^T \phi(\mathbf{x}_n), \beta^{-1}) \\ &= \frac{N}{2} \ln \beta - \frac{N}{2} \ln(2\pi) - \beta E_D(\mathbf{w}) \end{aligned} \quad (4.19)$$

where the sum-of-squares error function is defined by

$$E_D(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{t_n - \mathbf{w}^T \phi(\mathbf{x}_n)\}^2. \quad (4.20)$$

The maximum likelihood can be used to determine \mathbf{w} and β . First considered the maximisation with respect to \mathbf{w} , the maximisation of the likelihood function under a conditional Gaussian noise distribution for a linear model is equivalent to minimising a sum-of-squares error function given by $ED(\mathbf{w})$. The gradient of the log likelihood function 4.19 takes the form

$$\nabla \ln p(\mathbf{T}|\mathbf{w}, \beta) = \sum_{n=1}^N \{t_n - \mathbf{w}^T \phi(\mathbf{x}_n)\} \phi(\mathbf{x}_n)^T. \quad (4.21)$$

Setting this gradient to zero gives

$$0 = \sum_{n=1}^N t_n \phi(\mathbf{x}_n)^T - \mathbf{w}^T \left(\sum_{n=1}^N \phi(\mathbf{x}_n) \phi(\mathbf{x}_n)^T \right) \quad (4.22)$$

Solving for \mathbf{w} gives

$$\mathbf{w}_{ML} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{T} \quad (4.23)$$

which are known as the *normal equations* for the least squares problem. Here Φ is an $(N \times M)$ matrix, called the *design matrix*, whose elements are given by so that

$$\Phi = \begin{pmatrix} \phi_0(x_1) & \phi_1(x_1) & \dots & \phi_{M-1}(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \dots & \phi_{M-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \dots & \phi_{M-1}(x_N) \end{pmatrix}. \quad (4.24)$$

The quantity

$$\Phi^\dagger \equiv (\Phi^T \Phi)^{-1} \Phi^T \quad (4.25)$$

is known as *Moore-Penrose pseudo-inverse* of the matrix Φ (Rao (1971), Golub (1996)). It can be regarded as a generalisation of the notion of matrix inverse to nonsquare matrices. Indeed, if Φ is square and invertible, then using the property $(AB)^{-1} = A^{-1}B^{-1}$ gives $\Phi^\dagger \equiv \Phi^{-1}$.

Least Mean Squares Algorithm

Techniques, such as the maximum likelihood solution, which involve processing the entire training set in one go, can be computationally costly for large data sets. If the data set is sufficiently large, it may be worthwhile to use sequential algorithms, also known as on-line algorithms, in which the data points are considered one at a time, and the model parameters updated after each such presentation. Sequential learning is also appropriate for real time applications in which the data observations are arriving in a continuous stream, and predictions must be made before all of the data points are seen (Bishop, 2007).

The sequential learning algorithm can be obtained by applying the technique of stochastic gradient descent, also known as sequential gradient descent, as follows. If the error function

comprises a sum over data points $E = \sum_n E_n$, then after presentation of pattern n , the stochastic gradient descent algorithm updates the parameter vector \mathbf{w} using

$$\mathbf{w}^{(\tau+1)} = \mathbf{w}^\tau - \eta \nabla E_n \quad (4.26)$$

where τ denotes the iteration number, and η is a learning rate parameter. The value of η needs to be chosen with care to ensure that the algorithm converges. The value of \mathbf{w} is initialised to some starting vector $\mathbf{w}^{(0)}$. For the case of the sum-of-squares error function this gives

$$\mathbf{w}^{(\tau+1)} = \mathbf{w}^\tau - \eta(t_n - \mathbf{w}^{(\tau)T} \phi_n) \phi_n \quad (4.27)$$

where $\phi_n = \phi(x_n)$ (Bishop, 2007).

Regularised Least Squares

Another way of handling large data sets on line is to use regularised least squares. In order to control over-fitting a regularisation term to an error function can be introduced, so that the total error function to be minimised takes the form

$$E_D(\mathbf{w}) + \lambda E_W(\mathbf{w}) \quad (4.28)$$

where λ is the regularisation coefficient that controls the relative importance of the data-dependent error $E_D(\mathbf{w})$ and the regularisation term $E_W(\mathbf{w})$. One of the simplest forms of regulariser is given by the sum-of-squares of the weight vector elements

$$E_W(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w}. \quad (4.29)$$

Also considered the sum-of-squares error function given by

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{t_n - \mathbf{w}^T \phi(\mathbf{x}_n)\}^2 \quad (4.30)$$

then the total error function becomes

$$\frac{1}{2} \sum_{n=1}^N \{t_n - \mathbf{w}^T \phi(\mathbf{x}_n)\}^2 + \frac{\lambda}{2} \mathbf{w}^T \mathbf{w}. \quad (4.31)$$

This particular choice of regulariser is known in the machine learning literature as weight decay because in sequential learning algorithms, it encourages weight values to decay towards zero, unless supported by the data. In statistics, it provides an example of a parameter shrinkage method because it shrinks parameter values towards zero. It has the advantage that the error function remains a quadratic function of \mathbf{w} , and so its exact minimiser can be found in closed form. Specifically, setting the gradient of 4.31 with respect to \mathbf{w} to zero, and solving for \mathbf{w} as in section 4.3.2, this gives

$$\mathbf{w} = (\lambda \mathbf{I} + \Phi^T \Phi)^{-1} \Phi^T \mathbf{t}. \quad (4.32)$$

This represents a simple extension of the least-squares solution in 4.23 (Bishop, 2007).

4.3.3 Support Vector Machine

This section shall give a short introduction to the idea behind support vector machines. They are used in a variety of classification and regression applications and originally it was also planned to use one for the system to predict the finger positions. As it will become apparent, using support vector machine would be inappropriate and too time-consuming.

An important property of support vector machines is that the determination of the model parameters corresponds to a convex optimisation problem, and so any local solution is also a global optimum. Nevertheless, they suffer from a number of limitations. In particular, the outputs of support vector machines represent decisions rather than posterior probabilities. Also, the support vector machine was originally formulated for two classes, and the extension to more than two classes is problematic. There are some parameters that must be found using a hold-out method such as cross-validation. Finally, predictions are expressed as linear combinations of kernel functions that are centred on training data points and that are required to be positive definite (Bishop, 2007).

Other disadvantages is the limitation is speed and size, both in training and testing. Also the choice of the kernel is crucial (Burges, 1998).

4.4 Image Processing

In order to be able to efficiently interpret the ultrasound images, the relevant information has to be extracted. In this section, an introduction to image processing and a description of some image features is presented.

Nowadays, there is almost no area that is not impacted in some way by digital image processing. Images are produced by a variety of physical devices, including still and video cameras, scanners, radar, and ultrasound, and are used for a variety of purposes, including, medical, business, industrial, military, entertainment, and scientific. The interests in digital image processing stem from the improvement of pictorial information of human interpretation and the processing of scene data for autonomous machine perception. Digital image analysis is a process that converts a digital image into something other than a digital image, such as a set of measurement data or a decision (Shih, 2010).

In general, image processing operations can be categorised into four types:

1. Pixel operations: The output at a pixel depends only on the input at that pixel, independent of all other pixels in that image.
2. Local operations: The output at a pixel depends on the input values in a neighbourhood of that pixel. This operation can be adaptive because results depend on the particular pixel values encountered in each image region.
3. Geometric operations: The output at a pixel depends only on the input levels at some other pixels defined by geometric transformation. Geometric operations are different from global operations, such that the input is only from some specific pixels based on geometric transformation. They do not require the input from all the pixels to make its transformation

4. Global operations: The output at a pixel depends on all the pixels in an image. It may be independent of the pixel values in an image, or it may reflect statistics calculated for all the pixels, but not a local subset of pixels (Shih, 2010).

The system described in this thesis uses mainly local operations: The information of specific features, to be more precise, circular areas, is extracted.

4.4.1 Feature-specific Measurements

Measurements that can be performed on each of the individual features in images can be grouped into four classes: brightness, location, size, and shape. For each class, quite a variety of different specific measurements can be made, and there are also a variety of different ways to perform the operations. These measurements usually produce a numeric output suitable for statistical analysis, feature selection, or presentation graphics (Russ, 2006).

The ultrasound images are gray-value images and thus the information of the circular areas depends on its brightness as well as its orientation.

Brightness Measurements

The most typical range of brightness values for images is from 0 to 255 (8-bit range). Several such values can be combined to represent colour information. It is generally useful to be able to measure the mean intensity or a calibrated density value, as well as to find the brightest or darkest pixel values in each region or feature, and perhaps the standard deviation of the brightest values as a measure of variation or textures (Russ, 2006).

Determining Orientation

In several applications for measuring brightness values, the location of features is also needed for interpretation of the results. For a typical irregular feature extending over several pixels, there can be several different definitions of location, some easier to calculate than others.

The centroid location for instance, also called the centre of gravity, can be calculated correctly from a boundary representation such as chain code. For some purposes, the pixel brightness makes some pixels more important than others.

Closely related to the location of the centroid of a feature is the idea of determining its orientation. There are a number of different parameters that are used, including the orientation of the longest dimension in the feature and the orientation of the major axis on an ellipse fitted to the feature boundary. But just as the centroid is a more robust descriptor of the feature's location than is the midpoint, an orientation defined by all of the pixels in the image is often better than any of these because it is less influenced by the presence or absence of a single pixel around the periphery (Russ, 2006).

4.4.2 Gray Value Surfaces

This section in detail describes the properties of the image features. For each of the features (each feature is represented by a small, circular area) the information of the gray-value moments is used to extract the desired information.

The moment axis of a feature is the line around which the feature, if it were cut from rigid, uniform cardboard, would have the lowest moment of rotation. It can also be described as the axis that best fits all of the pixels, in the sense that the sum of the squares of their individual distances from the axis is minimised. This is the same criterion that can be used to fit lines to data points when constructing graphs. It is possible to weight each pixel with some value, such as the density, instead of letting each one vote equally (Russ, 2006).

The *Halcon* operator *fit surface first order* calculates the gray value moments and the parameters of the approximation of the gray values by a first order surface. The calculation is done by minimising the distance between the gray values and the surface. A first order surface is described by the following formula 4.33:

$$image(r, c) = \alpha(r - r_{center}) + \beta(c - c_{center}) + \gamma \quad (4.33)$$

where r denotes the row and c denotes the column. Thus r_{center} and c_{center} are the center coordinates of intersection of the input region with the full image domain. By the minimisation process the parameters from Alpha, Beta and Gamma are calculated. For fitting the surface, a least squares line fitting algorithm is used (MVTec Software GmbH, 2010).

For a better understanding, the parameters alpha (α), beta (β) and gamma (γ) of a surface first order are illustrated in figure 4.13. It can be seen that alpha denotes the gradient along row direction and beta along column direction with respect to the original image plain. Gamma donates the offset.

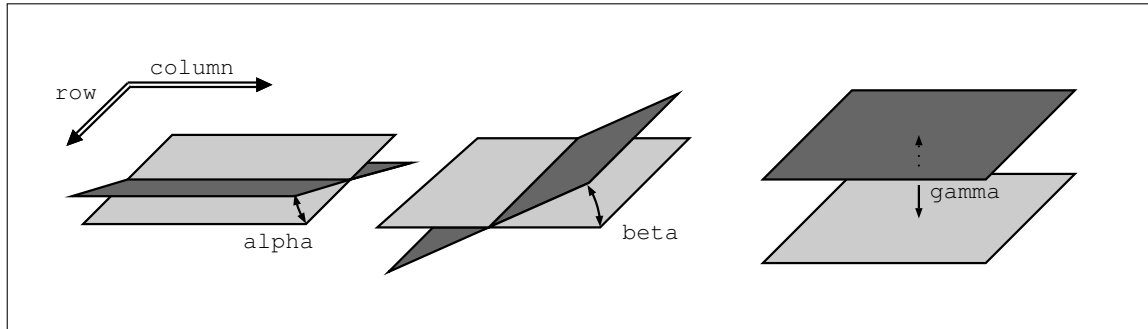


Figure 4.13: Parameters alpha, beta and gamma of a surface first order

4.4.3 Optical Flow

Another image processing tool called the optical flow is implemented to decrease the error during an experiment. A description of what the optical flow is exactly used for can be found in section 5.4.1.

The optical flow represents information about the movement between two consecutive images of a monocular image sequence. The movement in the images can be caused by objects that move in the world or by a movement of the camera (or both) between the acquisition of the two images. The projection of these 3D movements into the 2D image plane is called the optical flow (MVTec Software GmbH, 2010).

The two consecutive images of the image sequence are passed to the optical flow function. The computed optical flow is returned in a vector field. The vectors in the vector field represent the movement in the image plane between the two images. The point in image 2 that corresponds to the point (r, c) in image 1 is given by

$$(r', c') = (r + u(r, c), c + v(r, c)) \quad (4.34)$$

where $u(r, c)$ and $v(r, c)$ denote the value of the row and column components of the vector field at the point (r, c) (MVTec Software GmbH, 2010).

4.5 Phantom Limb Pain

Many patients with amputated limbs continue to have a vivid sensory experience of the missing limb, a phenomenon known as the phantom limb syndrome. The patient senses the presence of the missing limb, feels it move around, and even feels it try to shake hands when greeting someone. People often feel terrible pain in the phantom limb (Kandel, et al., 2000).

Since the system introduced in this thesis could be used for phantom limb pain treatment, this section describes causes, symptoms and treatments of phantom limb sensations.

The amputation of a limb is commonly followed by the sensation that the deafferented body part is still present. These non-painful phantom sensations may include a specific position, shape or movement of the phantom, feelings of warmth or cold, itching, tingling, or electric sensations. Pain in the body part that is no longer present occurs in 50-80 % of all amputees. Peripheral factors as well as a central pain memory might be causing phantom limb pain. The pain may be related to a certain position or movement of the phantom and may be elicited or exacerbated by a range of physical and psychological factors (e.g. changes in weather, pressure on the residual limb, emotional stress). It seems to be more intense in the distal portions of the phantom and can have several different qualities such as stabbing, throbbing, burning or cramping. Phantom limb pain is commonly confused with pain in the area adjacent to the amputated body part. Such pain is referred to as residual-limb or stump pain, and it is positively associated with phantom limb pain in most cases (Flor, 2002).

Phantom limb sensation and the pain associated with it have been attributed to impulses entering the spinal cord from the scar of nervous tissue in the stump. However, recent imaging studies by Vilayanur Ramachandran of the somatosensory cortex of patients who have lost a hand suggest another explanation for phantom limb sensations (Kandel, et al., 2000). Evidence was provided that these cortical changes are related less to referred sensations but have a close association with phantom limb pain. The larger the shifts, the greater the phantom limb pain (Flor, 2002).

4.5.1 Treatment of Phantom Limb Pain

Several studies, including large surveys of patients with amputations, have shown that most treatments for phantom limb pain are ineffective and do not take account of the mechanisms underlying the production of the pain. Table 4.8 lists commonly used treatments for phantom limb pain. The

use of a myoelectric prosthesis may influence phantom limb pain. The intensive use of myoelectric prosthesis was positively associated with both less phantom limb pain and less cortical reorganisation (Flor, 2002).

Results of studies by (Gagné, et al., 2009) also support the existence of a relationship between the ability to move the phantom limb and phantom limb pain.

Removing the scar or cutting the sensory nerves just above the stump does relieve pain in some cases (Kandel, et al., 2000).

Type of treatment	Examples
Pharmacological	Conventional analgesics, antidepressants, muscle relaxants
Surgical	Stump revision, dorsal column stimulation, deep brain stimulation
Anaesthetic	Nerve blocks, local anaesthesia
Psychological	Electromyographic biofeedback, temperature feedback, cognitive-behavioural pain management, hypnosis
Other	Acupuncture, physiotherapy, ultrasound, prosthesis training

Table 4.8: Commonly used treatments for phantom limb pain (Flor, 2002)

Ramachandran and Rogers-Ramachandran (1996) introduced another treatment approach to resurrect the phantom visually to study inter-sensory effects. In this procedure a mirror is placed vertically on the table so that the mirror reflection of the patient's intact hand is superimposed on the felt position of the phantom. The patient is viewing the reflection of the intact hand in the mirror, thus creating the illusion of observing two hands. If the patient sends motor commands to both arms to make mirror symmetric movements he will literally see his phantom hand resurrected and obeying his commands. Thus the patient receives positive visual feedback informing his brain that his phantom arm is moving correctly. Experiments show that this procedure seems to alleviate pain in some patients.

The system introduced in this thesis also delivers positive visual feedback to the patient performing an experiment. The patient sees a virtual hand on the screen moving the same way the absent hand would do. Therefore it is claimed that this system could be used to alleviate phantom limb pain.

5 Setup Specification

This section gives a detailed description of the setup's components. The main functions of the setup are grabbing the ultrasound images from the ultrasound system and, at the same time, recording finger positions gathered by the cyberglove. Furthermore all the data have to be processed which includes feature extraction, synchronisation and regression analysis. Additionally a virtual hand model is used to display both the system's output and the predefined stimuli movements (more about the stimuli movements in section 6). Figure 5.1 shows the outline of the setup. The data synchronisation and processing is enforced on a Windows PC equipped with a multi-core processor.

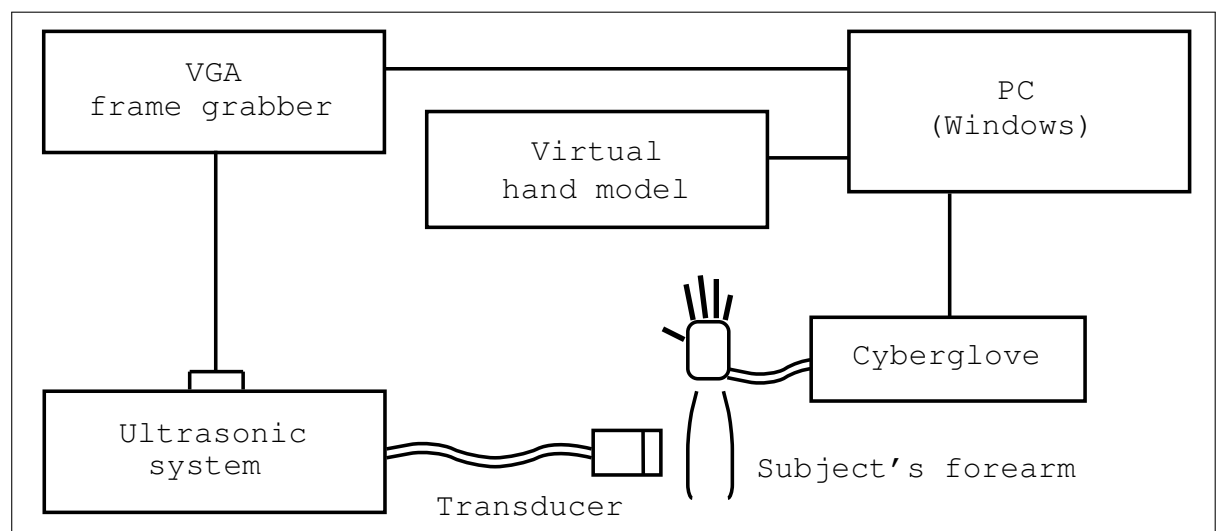


Figure 5.1: Principal outline of the setup

In principle, two different experiments with this setup are conducted. First, the preliminary experiments with healthy subjects to proof the feasibility of the approach, and secondly, the final experiments with an amputee. Even though the experiments are pretty similar, there is a difference of how to gather the ground truth of the finger positions. When conducting experiments with an amputee, who is obviously not able to wear the cyberglove on the amputated hand, there are two possibilities: either the subject wears a cyberglove on the healthy hand or the data from the virtual hand model is used as ground truth. In both cases the subject is forced to make synchronised finger movements with his two hands or with the virtual hand model.

5.1 Used Hardware

The setup basically consists of the following hardware:

- Ultrasound imaging system (generating ultrasound images)
- VGA frame grabber (grabbing ultrasound images)
- Cyberglove (gathering finger position data)
- Windows PC (data synchronisation and processing)

5.1.1 Ultrasound Imaging System

In figure 5.2 one can see the used *LOGIQ e* system from *General Electric Healthcare*. The system is portable and has the size of notebook.

Figure 5.3 shows the ultrasonic transducer of the type 12L-RS. The lens face (the actual part of the transducer in touch with patient's skin) is of size 45mm x 11 mm.

These kind of probes are usually used for the examination of small parts and peripheral, vascular structures. It has a wide field of view for B-Mode resolution and homogeneity (General Electric Co., 2006a).

The ultrasound system supports eight different application categories, including predefined settings for abdomen, obstetrics, gynaecology, cardiology, vascular, urology, and small Parts or paediatrics measurements (General Electric Co., 2006a).

In order to get the most meaningful and significant information possible, the settings for the ultrasound system have to be adjusted. Starting from the application category *vascular*, the main changing adjustments concern the image size, number of focus zones and depth. The optimal settings are listed in table 5.1.

The point in adjusting the setting for the ultrasound imaging system is to get images with significant information. The more structures visible in the ultrasound images, the better the results. It is not necessarily compelling that the ultrasound systems displays the correct anatomical correct boundaries between different tissue types. For example, tendons are not clearly visible, whereas the soft tissue enclosing tendons and muscles are clearly visible.

The most important thing is that the structures visible in the ultrasound images are moving coherently with the real finger movements. It does not really matter if the structures display tendons, muscles, soft tissue or something else.

As one can see in table 5.1, the imaging systems use two focal zones. It is important to put them into the area of interest to achieve a uniformly smooth image. Furthermore the depth is set to about four centimetre and the scan area was adjusted to the smallest reasonable size to maximise the frame rate. The gray map J is used because this map highlights tissue differentiation for a certain band of signals which is useful when more contrast for imaging structures is needed , e.g., renal, tendon, vein etc. One should also be aware that there are interdependencies between different setting, for example between the map, gain, and dynamic range. For more detailed information of the settings see also General Electric Co. (2006a).



Figure 5.2: Ultrasound system: *General Electric LOGIQ e* system



Figure 5.3: Ultrasonic transducer 12L-RS

Task	Explanation	Settings
Mode	Choose between M-Mode, B-Mode, Colour Flow Mode and Doppler Mode	B-Mode
Transducer Frequency (Frq)	Adjust transducer's frequency	12.0 MHz
Gain (GN)	Controls the amount of echo information displayed in an image	18
Edge Enhancement (E/A)	Brings out subtle tissue differences and boundaries by enhancing the gray scale differences corresponding to the edges of structures	5/4
Number of focus zones	Focal zone to tighten up the beam for specific areas	2
Map	Determines how the echo intensity levels received are presented as shades of gray	J/0
Depth (D)	The field of view	4.0 cm
Dynamic Range (DR)	Controls how echo intensities are converted to shades of gray, thereby increasing the adjustable range of contrast	93
Rejection	Selects a level below which echoes will not be amplified	0
Auto Optimise (AO)	Optimise the image based upon a specified region of interest or anatomy	100%
Time Gain Compensation (TGC)	Amplifies the returning signals to correct for the attenuation caused by tissues at increasing depth	low
Frame Rate (FR)	Resulting frame rate	20 Hz

Table 5.1: Optimised setting for the ultrasound system (see also General Electric Co., 2006b)

Finding the optimal settings is an empirical procedure. Settings may also be different from subject to another and therefore can be readjusted. However, in the experiments described in this thesis, the settings of the ultrasound imaging system have not been changed.

Transducer Position

As already mentioned in section 4.2.2, when discussing the anatomical cross sections, basically two different positions of where to place the transducer are described:

- Transducer position 1: distal, near the wrist (see figure 4.7, cross section A)
- Transducer position 2: mid-arm (see figure 4.7, cross section B)

In the preliminary experiments, mainly the transducer position 1 is used. Because in the beginning of implementing the setup, this position seemed to best capture tendons from all the muscles of the forearm. Still, the focus is on the flexor muscles and tendons, but also some extensor tendons are visible near the wrist.

Later on, when the feasibility of the approach was proven, also transducer position 2 was used. The advantage of the second position lies therein that it is easier to press the transducer onto the subject's skin. This position is also more comfortable for the subject, since there is more muscle tissue and the forearm is wider. Another difference are the muscles in the area of interest. In position 2, mainly the flexor muscles *digitorum profundus* and *digitorum superficialis* are examined. In comparison to transducer position 1, the imaging depth has to be increased to visualise also the extensor muscles. This however can decrease both the ultrasound system frame rate and the image processing speed.

Considering also that there exist different kind of hand amputations, it has to be proven that transducer position can be chosen freely. Results will show that it does not really matter where exactly the probe is placed as long as the desired muscles and tendons are in the ultrasound system's area of interest. This is especially a crucial part when conducting the experiments with an amputee.

5.1.2 VGA Framegrabber

Since the ultrasound system is not capable of delivering a stream of B-mode images directly to the PC, the ultrasound frames are grabbed from standard *Video Graphics Array* (VGA) interface using a frame grabber. The used VGA to Ethernet frame grabber by *Epiphan Systems Inc.* can be seen in figure 5.4.

Using the optimised settings, the ultrasound system generates images at a rate of 20Hz (see figure 5.1). These images are sent to the VGA interface at a resolution of 1024 x 768 at 60Hz.

The frame rate at which the grabber can grab the images, depends on the resolution of the VGA source, on the capacity frame grabber device, on the throughput of the Ethernet, and on the video capture workstation hardware. Because the frame grabber device includes hardware compression the maximum real frame rate also depends on the entropy of the input VGA signal. In other words, the frame rate depends on how well the signal can be compressed, or on how

much of the next frame is different from the previous one. The frame rate can range, from less than non compressing in the worst-case scenario to 125 frames per second (hard-wired limit) if almost nothing changes in the image (Epiphan Systems Inc., 2009).

The frame grabber also enables hardware cropping. Since the ultrasound systems delivers full screen images (see figure 5.2, only the necessary part (the actual ultrasound image) is grabbed.)



Figure 5.4: VGA to Ethernet frame grabber by *Epiphan Systems Inc.*

However, when the frame grabber tries to grab images at a certain rate, unwanted effects occur. Some of the new grabbed frames are identical with the previous ones. This is mainly because the grabbing rate is higher than the frame rate of the ultrasound system. In other words, the ultrasound system has not created a new ultrasound frame yet (with a frame rate of 20Hz) but has refreshed the VGA output (with a rate of about 60Hz). The only change in the image is a noise induced by the grabber.

Some of the grabbed frames also show a visual artifact where one part of the image is new but the other part of the image corresponds the previous frame. This effect is called *frame tearing* and is due to unsynchronised refresh rates of the grabber and the ultrasound system.

In order to detect these unwanted images, a simple algorithm using basic image processing functions has been implemented to make sure that no two identical or tear frames are saved or further processed. The algorithm can detect three kind of grabbed frames:

- Case 1: a completely new frame (a new frame created by the ultrasound system with noise added from the frame grabber).
- Case 2: a frame identical with the previous one (only the frame grabber noise changed the frame).
- Case 3: a tear frame (One part of the frame is new (created by the ultrasound image), the other part remains still from the previous frame).

In case 2 and case 3 the frames are invalid and not to be used.

As a first step, the difference of the gray values of two consecutive images is computed. To be more precise, the HALCON function called *absolute_diff_image* calculates the absolute difference between two images. The gray values g' of the output image are calculated from the gray values $(g1, g2)$ of the two input images as follows (MVTec Software GmbH, 2010):

$$g' = |(g1 - g2)| \quad (5.1)$$

Afterwards, the operator *intensity* calculates the mean and the standard deviation of the gray values in the image within Regions. If R is a region, p a pixel from R with the gray value $g(p)$ and F the plane ($F = |R|$), the features are defined by (MVTec Software GmbH, 2010):

$$Mean = \frac{1}{F} \sum_{p \in R} g(p) \quad (5.2)$$

$$Deviation = \sqrt{\frac{1}{F} \sum_{p \in R} (g(p) - Mean)^2} \quad (5.3)$$

The figure 5.5 shows the standard deviation values for a series of about 2800 consecutive frames. As one can see, it is possible to distinguish frames with frame grabber noise (case 2) from completely new frames (case 1). The deviation value of frames which are identical with the previous one is about 2.0, therefore a frame with a higher deviation value can be considered as new.

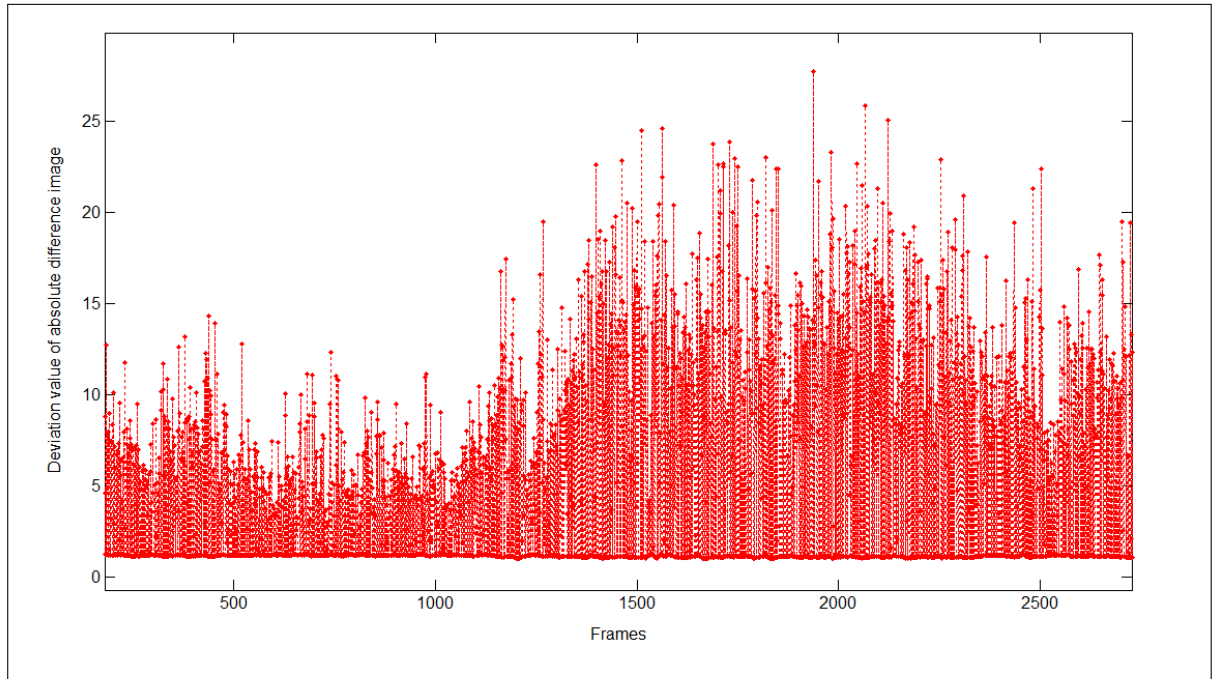


Figure 5.5: Tear frame detection

A close up of figure 5.5 is presented in figure 5.5. The plot shows about 18 frames; the 8 new frames can easily be distinguished from the 10 identical frames.

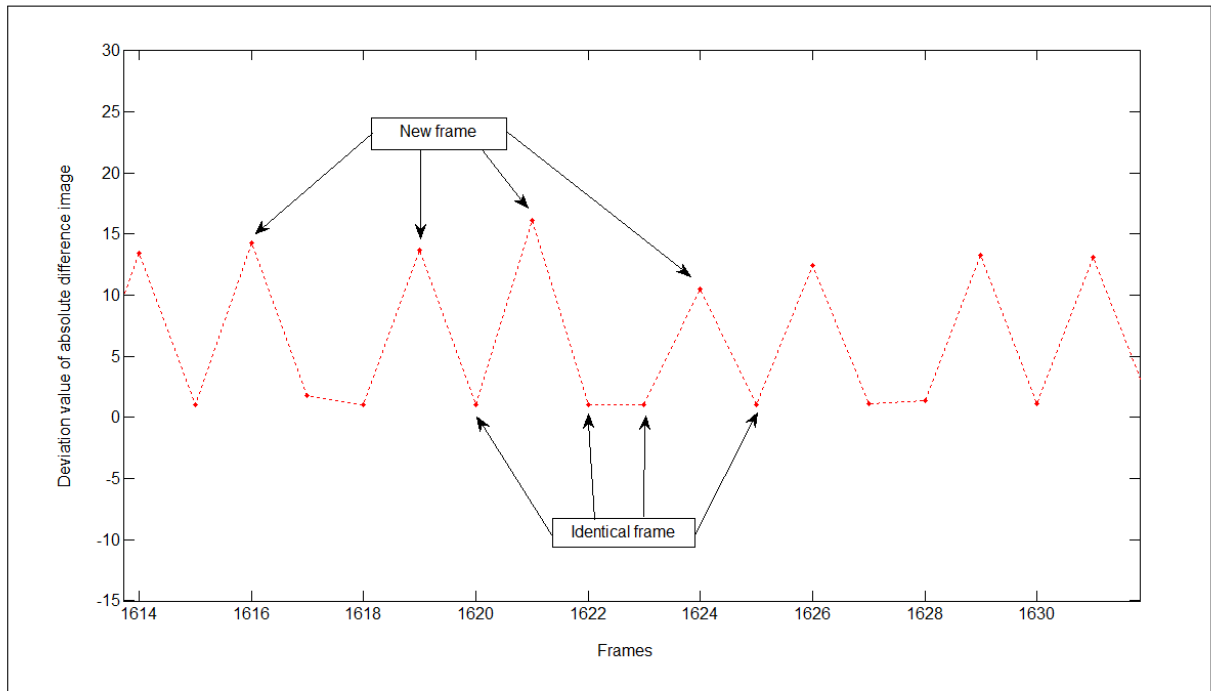


Figure 5.6: A close up of tear frame detection

In order to be able to detect tear frames, each frame is divided into two halves with corresponding regions. For each of the two regions the image processing steps are performed and lead to a slightly more sophisticated detection algorithm:

- Case a: Both regions detect a new frame: frame is valid.
- Case b: Both regions detect an identical frame: frame is invalid.
- Case c: One region detects a new frame and the other regions detects an identical frame: this frame is invalid, the next frame will be considered as valid.

5.1.3 Cyberglove

For the purpose of predicting the finger positions in the training phase it is essential to know the actual position of the fingers in the testing phase. The finger positions, or, to be more precise, the degree of flexion of the fingers are gathered with an 18-sensor right-handed glove by *CyberGlove Systems*. The glove features two bend sensors on each finger, four abduction sensors, plus sensors measuring thumb crossover, palm arch, wrist flexion, and wrist abduction. Figure 5.7 shows the position of the 18 sensors. The sensors exhibit a resistance which is proportionally related to the angles between pairs of hand joints. The manufacturer assures a resolution of less than one degree. For practical reasons the a cotton glove is worn below the Cyberglove.

In all the experiments described in this thesis, only 6 of the 18 sensors are used. These six degrees of motion are captured by considering the five metacarpophalangeal glove sensors, placed where the proximal phalanxes of the fingers meet the palm, plus the thumb/index abduction sensor for the thumb abduction/adduction. As also displayed in figure 5.7, sensors 13, 10, 7, 5 are

used for pinkie, ring, middle, and index flexion/extension. Sensor 4 is used for thumb adduction and sensor 1 for thumb rotation.

A careful hardware calibration enables obtaining a resolution of about 7:5 bits over the considered range, actually way below one degree in all cases. Values are normalised between 0 and 1 so that 0 corresponds to the fully extended position and 1 to the fully flexed position. The six motion values are streamed to a PC at the maximum rate allowed by the glove's underlying serial port connection, namely 88Hz.

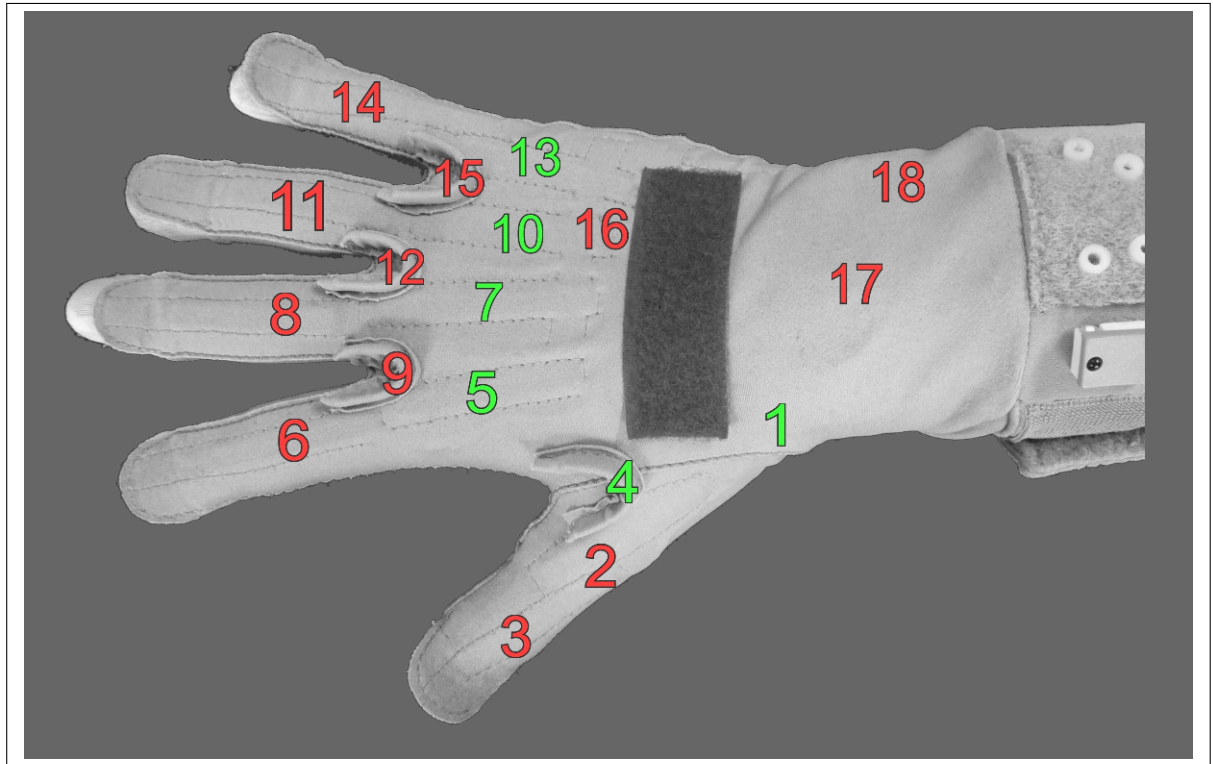


Figure 5.7: Cyberglove with 18 sensors. Only the six sensors 1,4,5,7,10 and 13 are used

In the final experiments conducted with an amputee, a left-handed glove was used instead of a right-handed one but there is no difference between the two gloves.

5.2 Used Software

The following software is used to synchronise and process all the data gathered from the peripheral devices:

- MATLAB: Used for data and regression analysis.
- Halcon (MVTec Software): The functions from the image processing library are mainly used for feature extraction.

- Microsoft Visual Studio C++: The setup's centrepiece gathers data from all peripheral devices and accurately timestamps each data unit.
- Blender: An already existing virtual hand model developed in Blender is used to display realistic and life-like finger movements.

A detailed description of the data processing procedure is given in the following section 5.3.

5.2.1 Virtual Hand Model

The virtual hand model is used to display the predicted position values as well as to demonstrate finger movements the subjects have to follow. Even though the hand model has 22 degrees of freedom, only six are used, namely the four metacarpophalangeal joints as well as the joints for thumb adduction and abduction (In figure 5.7 the corresponding sensor of the cyberglove are pictured).

Since only one degree of freedom is used, the position of the intermediate and distal phalanges is depending solely on the metacarpophalangeal joint value. For a better understanding, the three pinkie's joints are displayed in figure 5.8, namely the metacarpophalangeal joint (MCP), the proximal interphalangeal joint (PIP) and the distal interphalangeal joint (DIP). In order to imitate the human hand and generate realistic and life-like finger movements, the value of PIP is equivalent to the squared value of MCP. For example, if the MCP joint is flexed to 50%, the PIP joint is only flexed to 25%. The value of DIP is permanently set to zero.

5.3 Principle Procedure

This section describes in detail how the system processes the data and computes the output. Even though differences among the setups used for the different experiments exist, they all consist of three major phases, namely data gathering, training phase, and testing phase (see figure 5.9).

In the first phase, the data gathering phase, all the data needed for regression analysis are recorded (ultrasound images and finger positions). The finger position data are either gathered by the cyberglove or the virtual hand model. In this phase, the image features are directly extracted from the ultrasound images.

In the second phase the actual training takes place. As a first step, the gathered data have to be synchronised and normalised. Afterwards, the regression function is estimated.

In the third phase, the testing phase, the regression functions predicts the finger positions on the basis of new image features. The last step is to analyse the performance of the system.

There are two basic variations of the setup. An off-line version, essentially used for preliminary experiments, and an on-line version used for the final experiments. Figure 5.10 displays the off-line version and figure 5.11 the on-line version of the setup.

In the off-line version's data gathering phase, ultrasound images are recorded and the image features are extracted. At the same time, the finger positions are estimated, either using the cyberglove or the virtual hand model. The next step is to synchronise the image features and

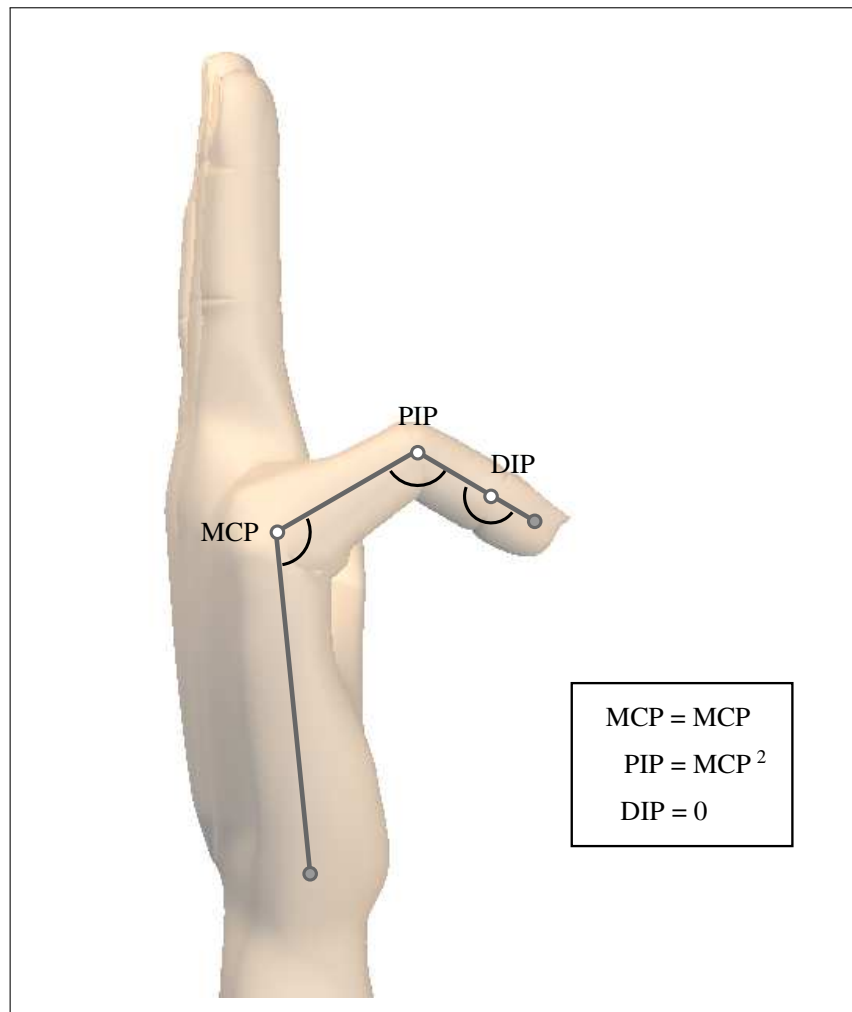


Figure 5.8: Flexion of the pinkie demonstrated using one degree of freedom

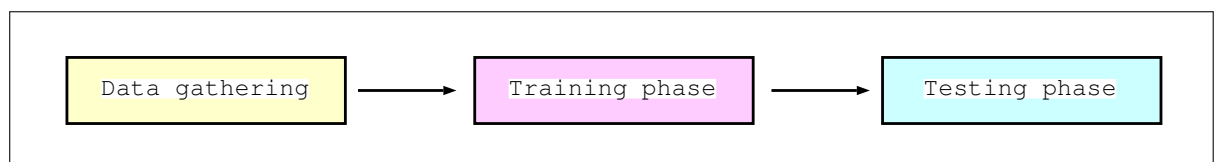


Figure 5.9: Three phases of data gathering, training phase, and testing phase

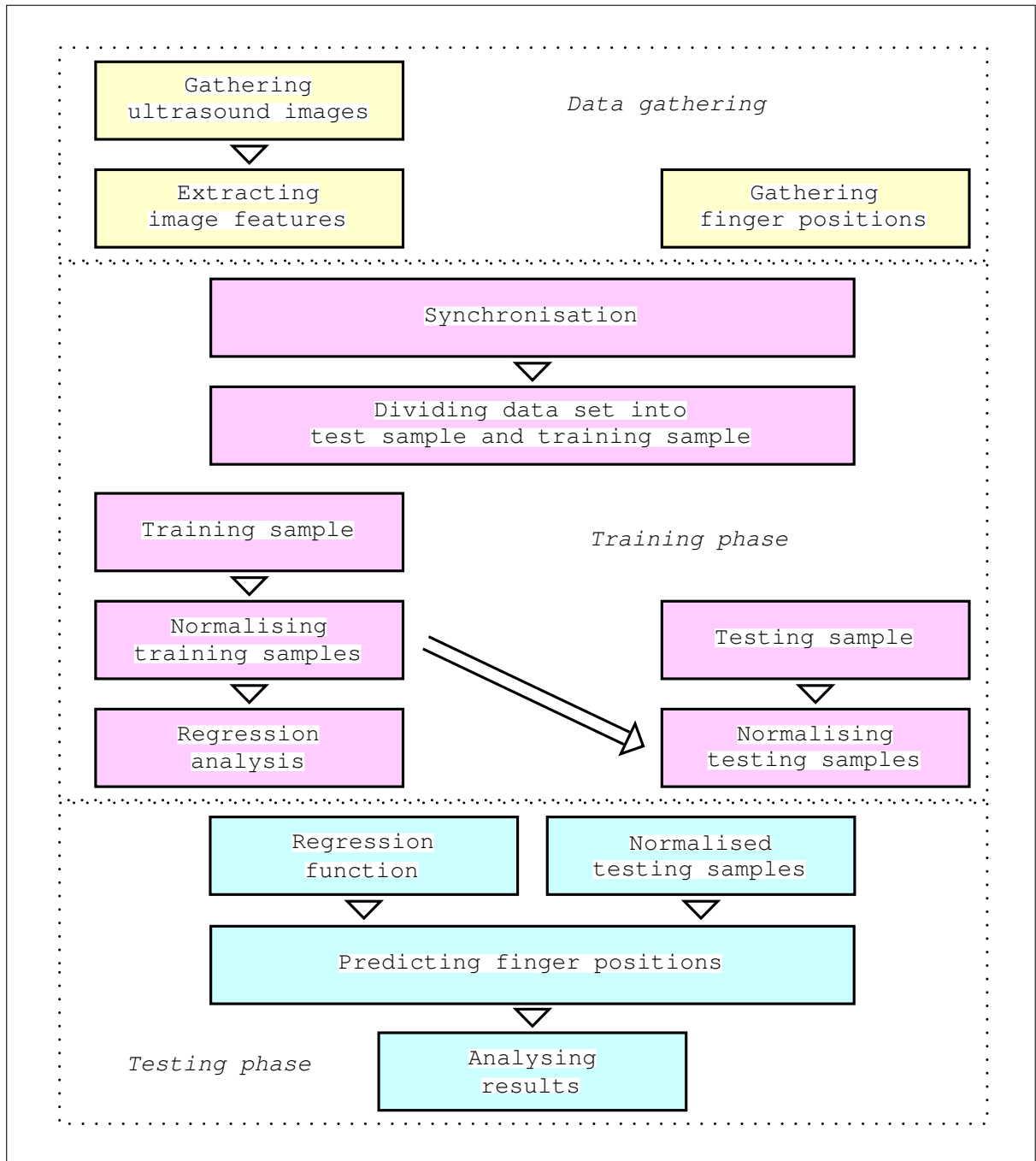


Figure 5.10: Outline of the off-line setup

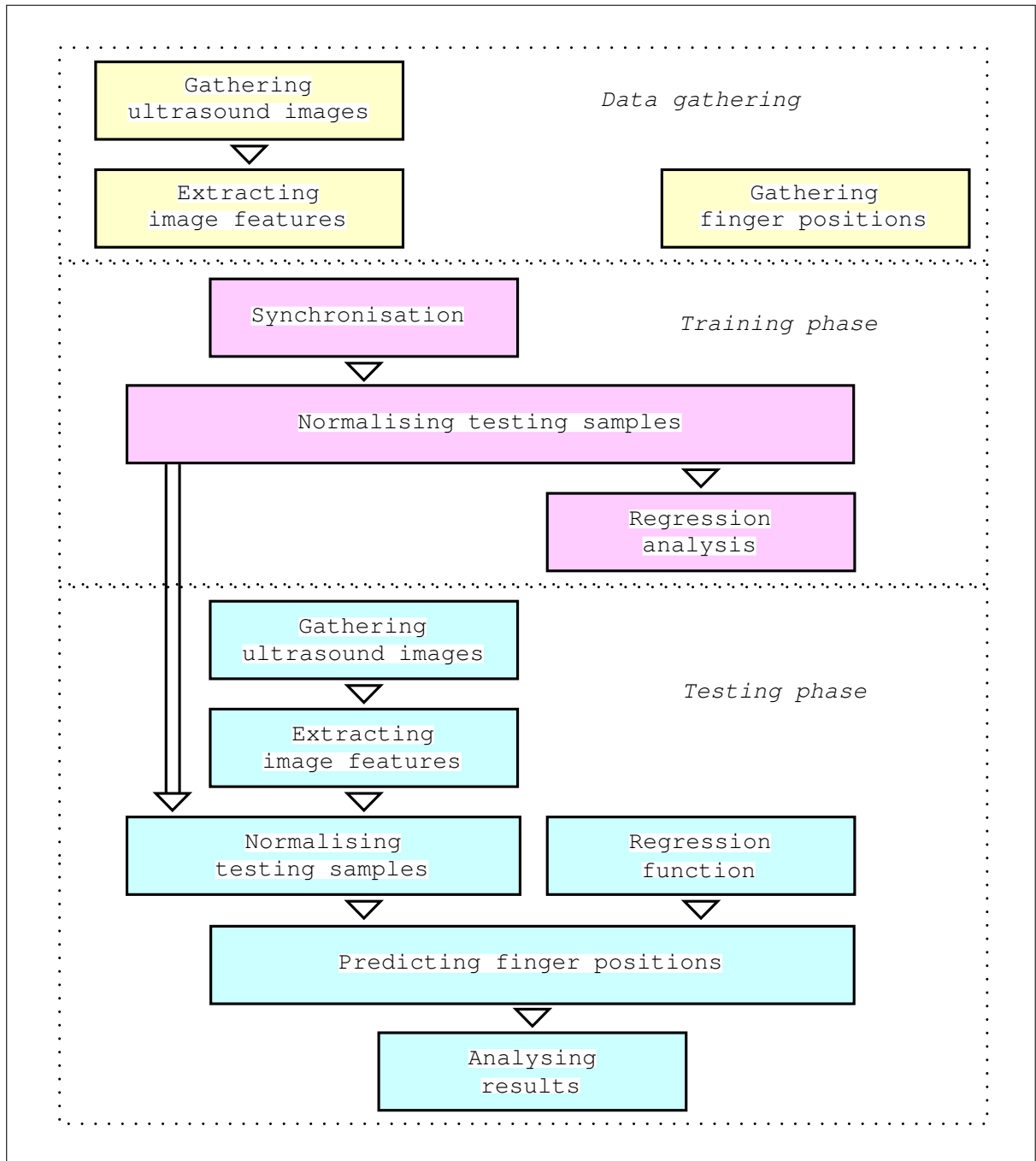


Figure 5.11: Outline of the on-line setup

finger positions with the help of each sample's timestamps.

Once the data are synchronised, the data set can be divided into testing and training set. There are several possibilities to divide the data set: In one case the training set consists of the first half and the testing set on the second half of the data set. It is also possible to train the set on every second sample and test is on the remaining samples. In both cases the samples can be shuffled beforehand. But of course also cross validation, as described in section 4.3.1 is way of examining how well the system generalises.

In order to avoid that some image features with large values have a stronger influence on the regression function, the next step is to normalise the training set data. This is done by subtracting the mean value and dividing by the standard deviation. With these very same statistics the testing samples are normalised as well.

The centrepiece of phase two is the regression analysis. The target is to find a function that predicts the finger positions in the best possible way. Therefore the unknown parameters of the function have to be estimated upon the training set. Finding the right kind of regression function can be a very crucial part. But the very first experiments revealed that there is a straightforward, linear relationship between the image features and the finger positions. This circumstance opens the possibility to build an on-line version of the setup, because both the estimation of the regression function and the prediction of the finger positions is very fast. A detailed description of estimating the regression function is presented in section 5.3.2.

The last step in the off-line version of the setup is to analyse the performance of the system and, if necessary, make modifications. The off-line version is also used to carry out a variety of different tests. For example finding the right number of image features and the average amount of training samples needed for the system in order to generalise well.

The on-line version of the setup, mainly used for the final experiments, is able to predict the finger positions on the fly. This means that the system, after it has been trained, estimates the positions values directly and displays the result on the screen via the virtual hand model.

In other words, the data gathered in phase one are not divided into a training and a testing set. Instead the training set consists of all the samples. Therefore the testing samples have to be gathered after the regression function has been built.

5.3.1 Data Gathering

In the first phase, both the ultrasound images of the forearm and the finger position data are gathered. Furthermore the image features are extracted. In this section the image feature properties as well as the different ways to gather finger positions will be discussed in detail.

Gathering Ultrasound Images

The ultrasound imaging system delivers a stream of B-mode images. These images are grabbed by the VGA frame grabber and are further processed. The first step therefore is to crop the image to the desired size. When the transducer is placed near the wrist (position 1) the size of the image usually is 508×210 pixel. Figure 5.12 displays such an ultrasound image.

Even though the ultrasound machine delivers gray images, these images are in the strict sense colour images (RGB colour space). This means that they consist three different colour channels for

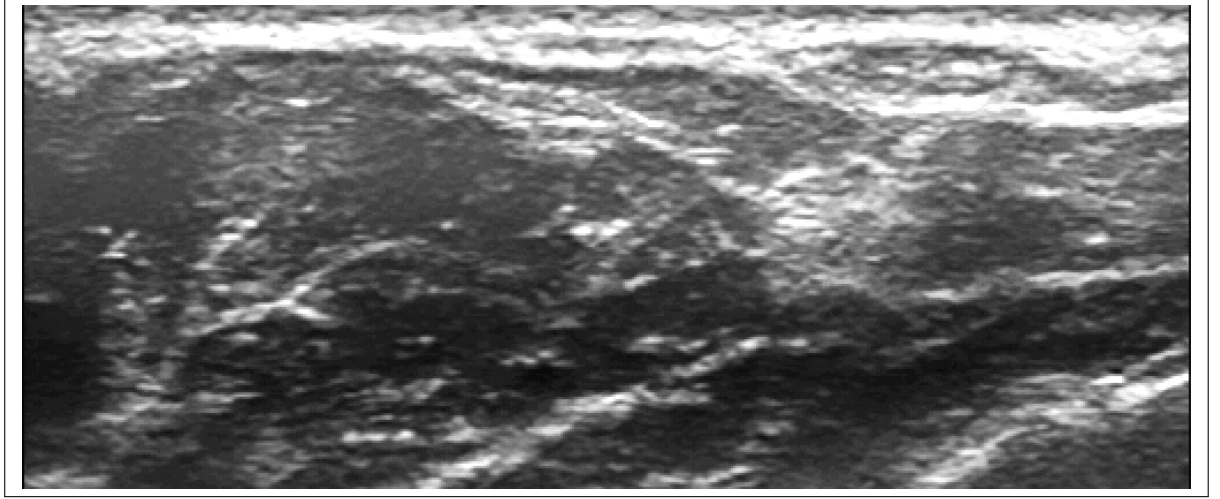


Figure 5.12: Example of a grayscale ultrasound image

each pixel, namely red, green and blue (Russ, 2006). As the images contain redundant information, the images are converted into real grayscale images at first. The easiest way to do so is to sum up the values of each channel and divide the sum by the number of channels, in this case three:

$$GrayImage(x, y) = \frac{R_{ColourImage}(x, y) + G_{ColourImage}(x, y) + B_{ColourImage}(x, y)}{3} \quad (5.4)$$

Feature Extraction

Once the images are converted, the next step is to extract the features. As already mentioned, the features are small circular first order surfaces. Each feature has a defined position, assigned by a equally spaced grid. Figure 5.13 shows the grid, in this case consisting of 67 points, displayed by crosses. The distance between to grid point is denoted by d . As one can see, each second row and each second column are shifted by $d/2$.

The size of the feature area is defined by the radius r (see figure 5.13). The figure also shows that the feature areas overlap each other. Thus the features, up to a certain extend, contain redundant information. But it is obvious that the degree of overlapping depends on the parameters d and r , which have a major impact on both the properties of the grid and the features. Thus the algorithm that generates the grid also has to make sure that no feature points are aligned too close to the border. Otherwise the feature areas jut out over the edge of the image and therefore deliver erroneous results.

Figure 5.15 shows what spots in the ultrasound image are covered by the features. Depending on the parameters d and r , some spots may not be covered.

How the parameters d and r influence the number of features can be seen in figure 5.16. The smaller the distance d , the higher the number of features. But also the radius influences the number of features. The bigger the radius r , the larger the distance from the features to the image border and the smaller the number of features. This impact is rather small but also

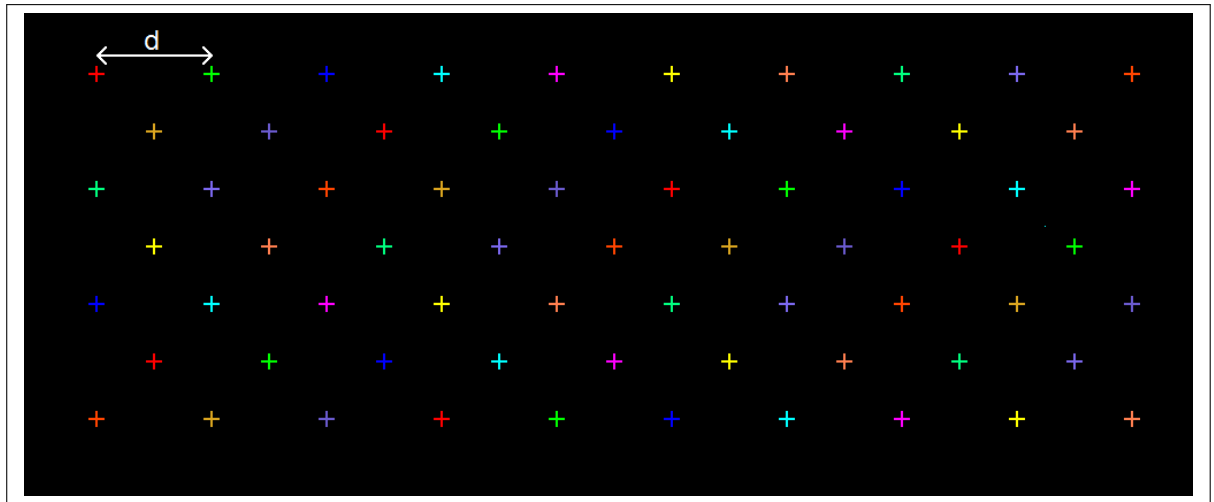


Figure 5.13: Grid points

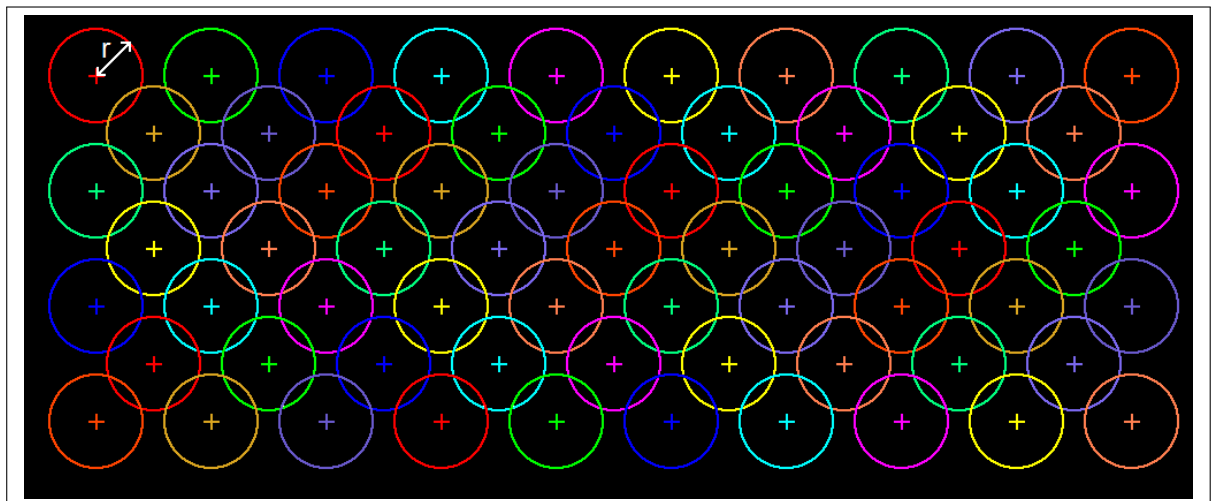


Figure 5.14: Grid points with circle areas

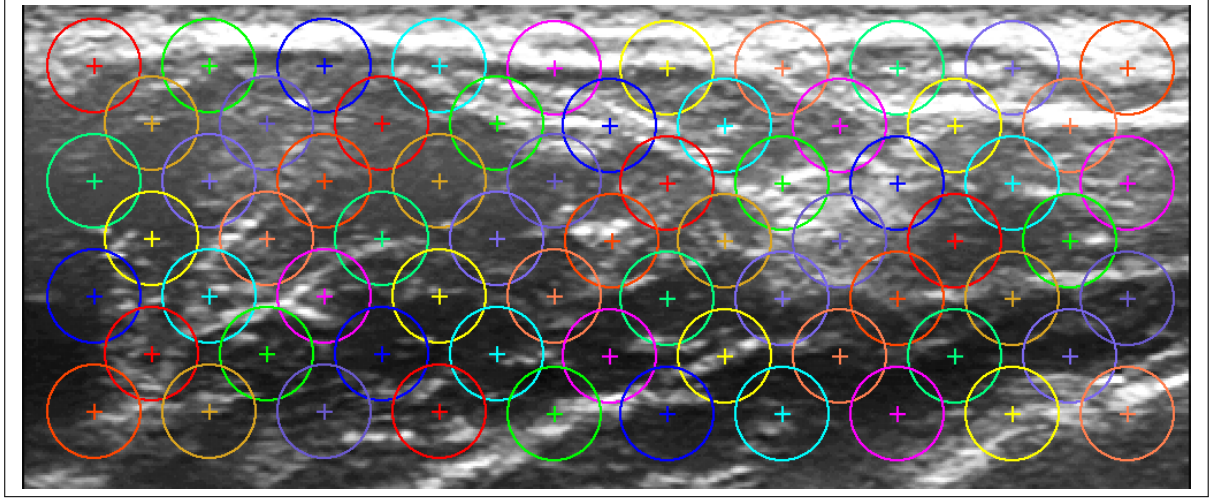


Figure 5.15: Feature points with corresponding areas

explains that the number of feature points increases not constantly as the radius decreases (can be seen in each row at the left hand side in figure 5.16).

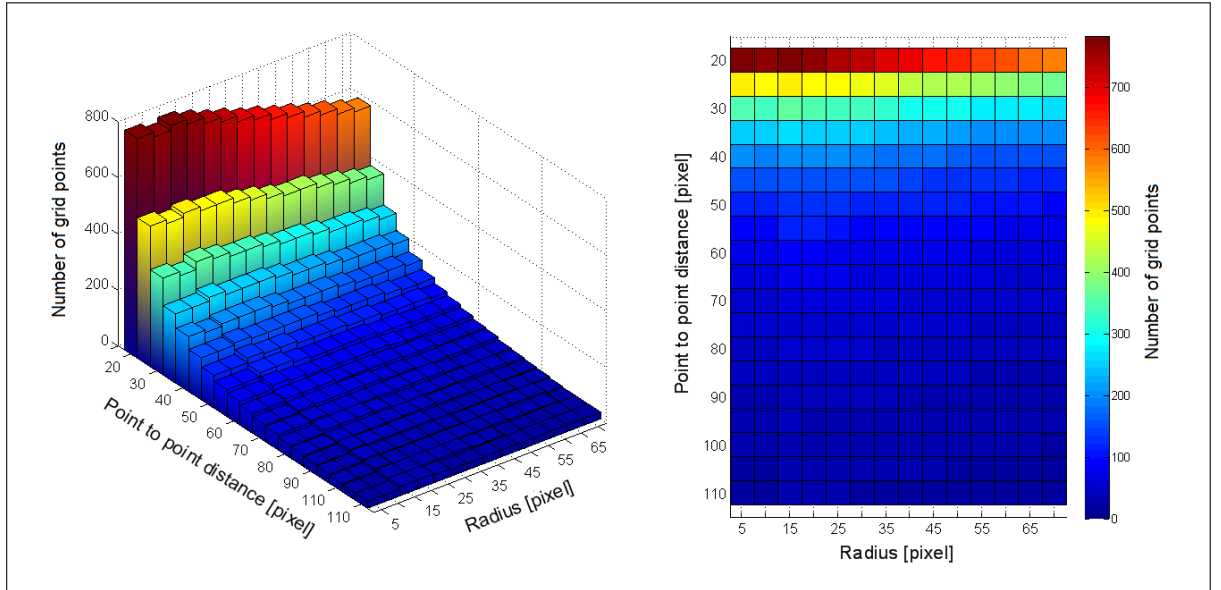


Figure 5.16: Number of grid points change when adjusting the radius and point to point distance

The number of grid points is critical because the more feature values, the higher to computation cost. The number of samples needed for training is also influenced by this number. There must be at least three times as many samples as feature points, because one feature point delivers three feature values (α , β and γ).

A grid consisting of f grid points results in $3f$ image feature values. These extracted feature values can be denoted by a vector. Extracting the features for a sequence of images i leads to i

vectors that can be denoted by a matrix. As already mentioned, each image, and therefore also each feature vector, owns a specific timestamp. This is very important for the synchronisation of image features with finger positions. The dimensions of the matrix for raw image features are therefore $i \times 3f$:

$$M_{RIF} = \begin{bmatrix} \alpha_{1,1} & \beta_{1,1} & \gamma_{1,1} & \alpha_{1,2} & \beta_{1,2} & \gamma_{1,2} & \dots & \alpha_{1,f} & \beta_{1,f} & \gamma_{1,f} \\ \alpha_{2,1} & \beta_{2,1} & \gamma_{2,1} & \alpha_{2,2} & \beta_{2,2} & \gamma_{2,2} & \dots & \alpha_{2,f} & \beta_{2,f} & \gamma_{2,f} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{i,1} & \beta_{i,1} & \gamma_{i,1} & \alpha_{i,2} & \beta_{i,2} & \gamma_{i,2} & \dots & \alpha_{i,f} & \beta_{i,f} & \gamma_{i,f} \end{bmatrix}. \quad (5.5)$$

Figure 5.17 displays a sequence of six different image features (one feature per row). Each feature area is displayed at then different times. The feature values alpha, beta and gamma itself are represented by little vectors with the starting point of the vector in the centre of the circular area. Alpha and beta denote the angle of the vector and gamma denotes its length. This figure illustrates how the feature values change with respect to the grayscale values of the ultrasound image.

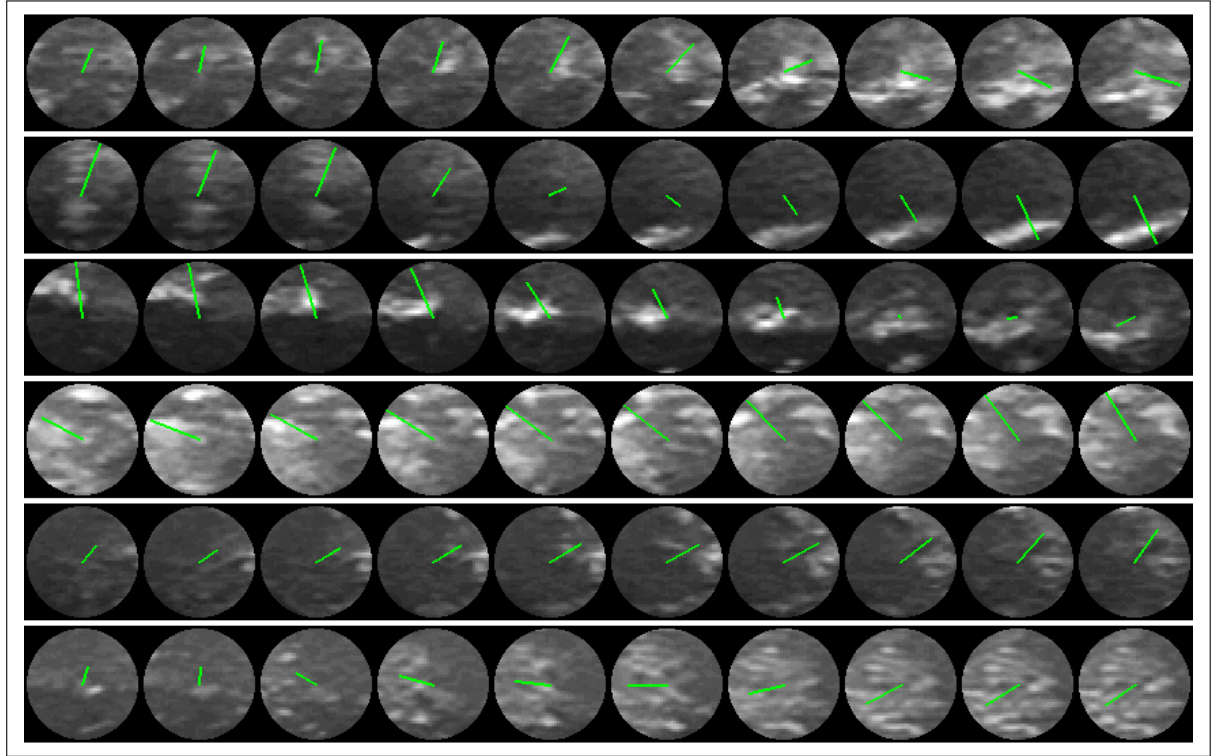


Figure 5.17: Sequence of six image features

Gathering Finger Positions

There exist two possibilities to gather information of the finger motion. Either use the glove presented in section 5.1.3, or use the virtual hand model. In the latter case, the lifelike hand model displays predefined finger movements and the subject has to imitate them. It is very

important that the virtual finger motions and the real finger motions are as synchronous as possible.

When using the cyberglove to gather the finger positions, healthy subjects wear the glove on the same arm as the ultrasound images are taken from. Amputees would wear the glove on the other hand, and just like when using the virtual hand model, try to make synchronous movements.

In all cases, both the glove and the hand model deliver six values, one for each degree of motion (for purposes of generality the number of degrees of motion is denoted by d). The values are normalised between 0 and 1 so that 0 corresponds to the relaxed position and 1 to the fully flexed position. Just like the ultrasound images, each sample owns a specific timestamp. The matrix for the finger position data therefore is of size $j \times d$, where j denotes the number of samples:

$$M_{FP} = \begin{bmatrix} DoM_{1,1} & DoM_{1,2} & \dots & DoM_{1,d} \\ DoM_{2,1} & DoM_{2,2} & \dots & DoM_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ DoM_{j,1} & DoM_{j,2} & \dots & DoM_{j,d} \end{bmatrix}. \quad (5.6)$$

The glove position values are streamed to a PC via serial port connection. The virtual hand model is easily controlled using the *user datagram protocol* (UDP). Thus the position values are the very same that also control the model.

5.3.2 Training Phase

In the second phase, the actual training phase, the data gathered are synchronised and normalised. After that, the regression function is evaluated.

Synchronisation

The synchronisation of finger positions and image feature is simply done by linear interpolation. Each sequence is interpolated over the timestamps of the image features. The image features stay the same.

Since the glove data are gathered at about 80 Hz and the ultrasound image frame rate is about 20 Hz, there are approximately four times more samples of finger positions. One can see that also in figure 5.18. The black line (with black dots) represents the original position values, the red line (with red circles) the new position values, interpolated over the timestamps of the image features. Figure 5.18 displays the positions for one degree of motion. The synchronisation has to be conducted for all six degrees of motion.

When the finger position matrix M_{FP} had the dimensions $j \times d$, the dimensions changed to $i \times d$. In other words, there are exactly as many finger positions samples as image feature samples. The aim of synchronisation is to have image feature samples and finger positions at the exactly the same instant of time. Since the goal is to predict the finger positions, the image features can be simply denoted by *samples* and the finger positions as *targets*.

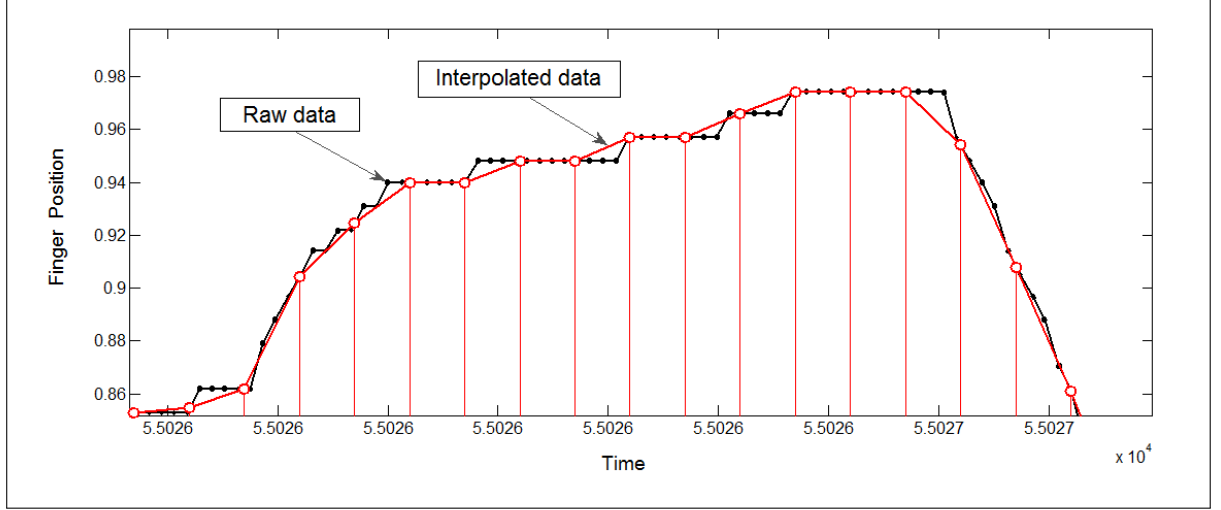


Figure 5.18: Interpolation of new finger position data

Normalisation of Image Features

Before the feature values are normalised, the data set has to be divided into a training set and a test set (the training set consists of training samples and training targets, the test set of test samples and test targets, respectively). It is also important to notice that the mean and standard deviation values are only obtained from the training samples. Obviously the statistics of the test samples are not known in advance. The test samples are normalised with the very same statistics of the training samples.

In case of the on-line version of the setup, the training set consists of all the data. Thus the mean and standard deviation values have to be saved for the testing phase.

The actual normalisation of the training samples has to be conducted for each of the $3f$ features. The new matrix for the training samples (image features) is denoted by

$$TrSample = \begin{bmatrix} \frac{\alpha_{(1,1)} - \hat{\alpha}_1}{\sigma_{\alpha_1}} & \frac{\beta_{(1,1)} - \hat{\beta}_1}{\sigma_{\beta_1}} & \dots & \frac{\gamma_{(1,f)} - \hat{\gamma}_f}{\sigma_{\gamma_f}} \\ \frac{\alpha_{(2,1)} - \hat{\alpha}_1}{\sigma_{\alpha_1}} & \frac{\beta_{(2,1)} - \hat{\beta}_1}{\sigma_{\beta_1}} & \dots & \frac{\gamma_{(2,f)} - \hat{\gamma}_f}{\sigma_{\gamma_f}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_{(i,1)} - \hat{\alpha}_1}{\sigma_{\alpha_1}} & \frac{\beta_{(i,1)} - \hat{\beta}_1}{\sigma_{\beta_1}} & \dots & \frac{\gamma_{(i,f)} - \hat{\gamma}_f}{\sigma_{\gamma_f}} \end{bmatrix} \quad (5.7)$$

that leads to

$$TrSample = \begin{bmatrix} \hat{\alpha}_{(1,1)} & \hat{\beta}_{(1,1)} & \hat{\gamma}_{(1,1)} & \hat{\alpha}_{(1,2)} & \hat{\beta}_{(1,2)} & \hat{\gamma}_{(1,2)} & \dots & \hat{\alpha}_{(1,f)} & \hat{\beta}_{(1,f)} & \hat{\gamma}_{(1,f)} \\ \hat{\alpha}_{(2,1)} & \hat{\beta}_{(2,1)} & \hat{\gamma}_{(2,1)} & \hat{\alpha}_{(2,2)} & \hat{\beta}_{(2,2)} & \hat{\gamma}_{(2,2)} & \dots & \hat{\alpha}_{(2,f)} & \hat{\beta}_{(2,f)} & \hat{\gamma}_{(2,f)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\alpha}_{(i,1)} & \hat{\beta}_{(i,1)} & \hat{\gamma}_{(i,1)} & \hat{\alpha}_{(i,2)} & \hat{\beta}_{(i,2)} & \hat{\gamma}_{(i,2)} & \dots & \hat{\alpha}_{(i,f)} & \hat{\beta}_{(i,f)} & \hat{\gamma}_{(i,f)} \end{bmatrix}. \quad (5.8)$$

The mean values can be denoted as vector with dimensions $1 \times 3f$

$$Mean_{tr} = (\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\alpha}_2, \dots, \hat{\gamma}_f) \quad (5.9)$$

and a vector containing all $3f$ standard deviation values:

$$Std_{tr} = (\sigma_{\alpha_1}, \sigma_{\beta_1}, \sigma_{\gamma_1}, \sigma_{\alpha_2}, \dots, \sigma_{\gamma_f}). \quad (5.10)$$

Since the values of the training and test targets (finger positions) all range between 0 and 1, no further normalisation is needed.

The new matrix containing the training samples is denoted by T_{tr} and has of course the same dimensions as the old matrix containing the raw values M_{RIF} .

Regression Analysis

The centrepiece of the finger prediction system is the regression analysis. As already stated before, the finger positions are linearly related to the image features, thus the regression function is also linear and basically consists of parameters, also called weights. The number of weights equals the number of image features. One can also imagine this regression function as a hyperplane with $3f$ dimensions.

The first step is create the design matrix M . It consists of the training samples and also includes constant terms, in this case a column of ones. The constant term has also be added to the training targets as well as test samples and test targets. The design matrix is denoted by

$$M = \begin{bmatrix} 1 & \hat{\alpha}_{(1,1)} & \hat{\beta}_{(1,1)} & \hat{\gamma}_{(1,1)} & \hat{\alpha}_{(1,2)} & \hat{\beta}_{(1,2)} & \hat{\gamma}_{(1,2)} & \dots & \hat{\alpha}_{(1,f)} & \hat{\beta}_{(1,f)} & \hat{\gamma}_{(1,f)} \\ 1 & \hat{\alpha}_{(2,1)} & \hat{\beta}_{(2,1)} & \hat{\gamma}_{(2,1)} & \hat{\alpha}_{(2,2)} & \hat{\beta}_{(2,2)} & \hat{\gamma}_{(2,2)} & \dots & \hat{\alpha}_{(2,f)} & \hat{\beta}_{(2,f)} & \hat{\gamma}_{(2,f)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \hat{\alpha}_{(i,1)} & \hat{\beta}_{(i,1)} & \hat{\gamma}_{(i,1)} & \hat{\alpha}_{(i,2)} & \hat{\beta}_{(i,2)} & \hat{\gamma}_{(i,2)} & \dots & \hat{\alpha}_{(i,f)} & \hat{\beta}_{(i,f)} & \hat{\gamma}_{(i,f)} \end{bmatrix} \quad (5.11)$$

and the training target by:

$$T_{tr} = \begin{bmatrix} 1 & DoM_{1,1} & DoM_{1,2} & \dots & DoM_{1,d} \\ 1 & DoM_{2,1} & DoM_{2,2} & \dots & DoM_{2,d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & DoM_{j,1} & DoM_{j,2} & \dots & DoM_{j,d} \end{bmatrix}. \quad (5.12)$$

As discussed in section 4.3.2, the parameters of the regression function are represented by the K can be estimated similar to formula 4.23 by

$$K = (M^\dagger T_{tr})^T \quad (5.13)$$

where M^\dagger again is the pseudo inverse of the design matrix M

$$M^\dagger \equiv (M^T M)^{-1} M^T. \quad (5.14)$$

Since the matrix M has the dimensions $i \times (3f + 1)$ (number of features times number of samples) the pseudo inverse matrix M^\dagger has dimensions $(3f + 1) \times i$. The dimensions of the training target matrix are $i \times d$ and from this it follows that the matrix K has dimensions $d \times (3f + 1)$. Each row represents a degree of motion. For example, with six degrees of motion and 67 feature points and $67 \times 3 = 201$ features, K 's dimensions are 6×202 .

It is very important that there at least exist as much training samples as features, otherwise there are insufficient data to estimate the least square model.

5.3.3 Testing Phase

In the testing phase, the regression function is tested on the test samples and the output is compared with the test targets.

Predicting Finger Positions

The first step again is to normalise the test samples (image features). It is basically done the same way the training samples are normalised. Instead of evaluating the mean and standard deviation from the test samples, the already computed values of $Mean_{tr}$ and Std_{tr} are used (under the assumption that the training and test samples are equally distributed).

Assuming that, in case of the on-line model, a new ultrasound image is captured, the features are extracted and normalised, the features can be denoted by the test sample vector S_{te} . As mentioned before, it is needed to add a constant term to the test samples. The predicted targets values can be estimated as follows

$$T_p = KS_{te} \quad (5.15)$$

Since the dimensions of the matrix K is $d \times (3f + 1)$ the sample vector S_{te} must have the dimensions $(3f + 1) \times 1$. It follows that the dimensions of the predicted target value vector are $d \times 1$. That way, with six degrees of motion, six predicted target values are obtained.

In case of the off-line setup, when not only one test sample is used, S_{te} is a matrix with dimensions $(3f + 1) \times i$ containing all the samples of the test set. In that case the predicted target matrix has dimensions $i \times 6$.

Using the on-line version, the predicted target values can be displayed directly on the screen via the virtual hand model.

Error Analysis

After the target values are predicted, two error measures and the correlation coefficient are evaluated. The mean absolute error (ERR) is evaluated as

$$ERR = \frac{1}{3f} \sum_{i=1}^{3f} |T_{p,i} - T_{te,i}| \quad (5.16)$$

and the normalised square-root mean-square error (NRMSE) as

$$NRMSE = \frac{\sqrt{\frac{1}{3f} \sum_{i=1}^{3f} (T_{p,i} - T_{te,i})^2}}{T_{te,max} - T_{te,min}}. \quad (5.17)$$

Just like presented in section, the correlation coefficient (CC) is evaluated as

$$CC[T_p, T_{te}] = \frac{cov[T_p, T_{te}]}{\sqrt{var[T_p]} \sqrt{var[T_{te}]}}. \quad (5.18)$$

5.4 Advanced Modifications

During the implementation of the setup, further modifications were realised in order to improve the performance of the system. Such modifications include the probe motion compensation as well as signal processing.

5.4.1 Probe Motion Compensation

In order for the system to work fine, the ultrasound probe must be still with respect to the forearm section of interest. If this is not the case the feature points move with respect to the image and the image structures. Thus the system is not able to predict the finger positions correctly.

Due to the fact that the transducer unconsciously drifts in time during one experiment, a compensation mechanism based upon the optical flow is implemented. The optical flow is performed from time to time on specific frames which will result in shifting the feature points in order for them to maintain the original position, relative to the initial frame.

In order to apply the probe motion compensation technique, it is essential to know the instant of time the subject's hand was in a predefined position, called resting position. In this position, all the fingers are stretched out and the angle between the thumb and index finger measures about 60 degrees. The resting position of the hand is shown in figure 6.1, hand motion 1.

At the beginning of an experiment the subject's hand is situated in this position and an ultrasound image is defined as initial frame. This image is used as the reference image. The images gathered during the experiment, when the subject's hand is situated in the resting position are then compared with the reference image and the offset is evaluated.

In the preliminary experiments described in section 6.1, this is only possible off line. During one session the hand acquires the resting position several times. In order to identify all images of the hand being in the resting position, the sequence of the virtual hand model that displays the finger movements to be mimicked, is synchronised with the images. Thus each image can be labelled with a corresponding movement value, and moreover identify all images of the hand situated in the resting position.

The probe motion compensation is basically based upon the optical flow. In order to apply this technique on two images, better results can be achieved when the images are smoothed beforehand. The Halcon function `smooth_image` smooths grayscale images using recursive filters

(see also MVTec Software GmbH, 2010). Image 5.19 displays an example of such a smoothed ultrasound image.

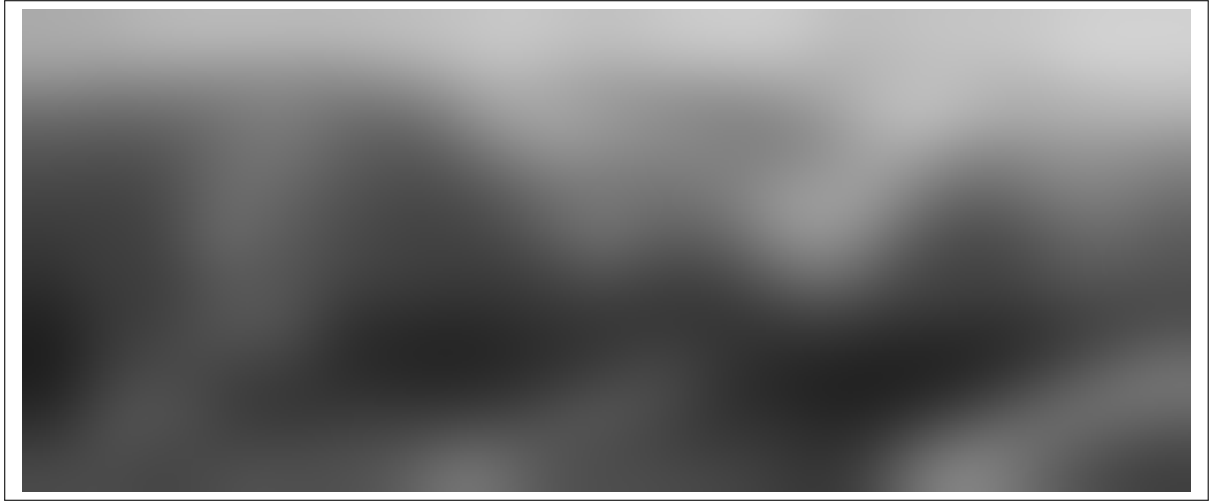


Figure 5.19: Smoothed ultrasound image before applying the optical flow algorithm

The next step is to apply the optical flow function and make the information about the movement available for further processing. The Halcon function *vector_field_to_real* converts a vector field (resulting from the optical flow) into two real-valued images. The output images contain the vector components in the row and column direction, respectively (MVTec Software GmbH, 2010). Image 5.20 displays the vector field generated by optical flow function. As one can see, the vectors mostly point into the right, or top-right positions. This indicates that the second image was moved to the right.

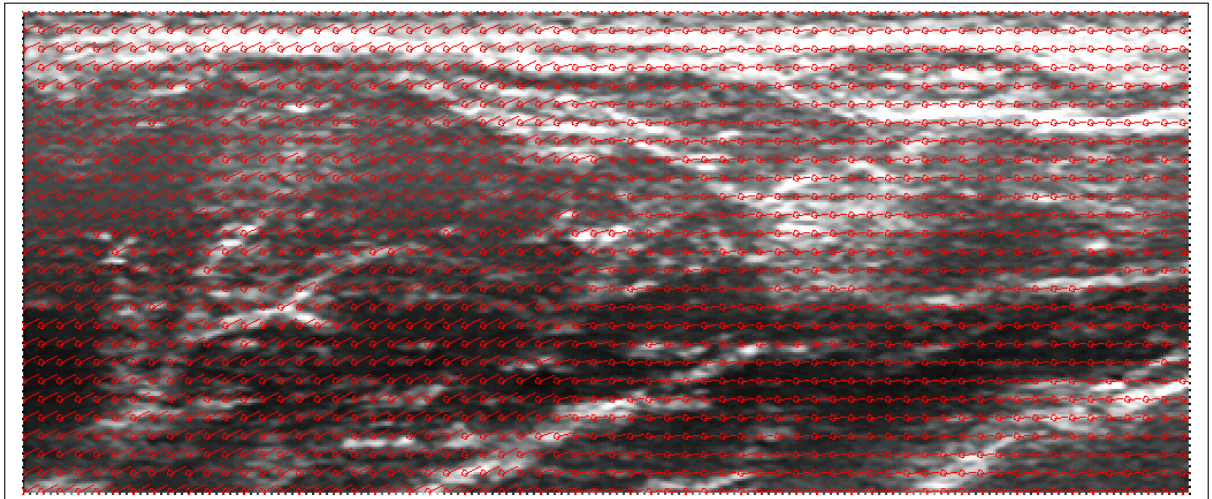


Figure 5.20: Vector components obtained after applying the optical flow algorithm

Figure 5.21 displays a close up of the vector components (the circles indicate the original positions). The vector components can be used to shift the feature points according to the

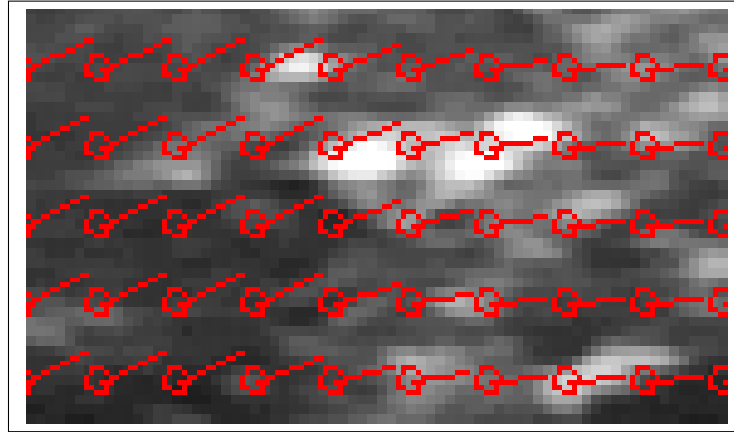


Figure 5.21: A close of the vector components obtained after applying the optical flow algorithm

movement between the two ultrasound images. The result of the probe motion compensation technique can be seen in figure 5.22. It is important to be aware that the vector field does not need to be homogeneous. To be more precise, all the vectors do not have to point into the very same direction. Different vector components for different regions may emerge mostly due to rotary movements induced by the subject's arm.

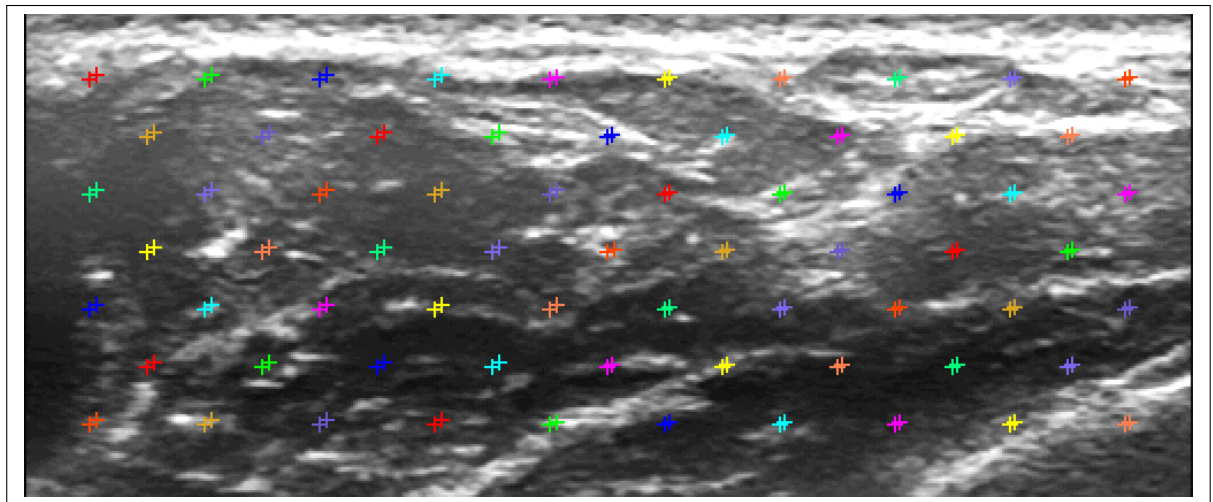


Figure 5.22: Shifting feature points

The parameters of the optical flow function, such as smoothness factor or algorithm technique, influence both the compensation time and the resulting vector components. Since the images differ in between subjects, these parameters may be adjusted to obtain optimal results (for information also see MVTec Software GmbH, 2010).

Compensation Off-line

When analysing the data of the preliminary experiments, the first thing is to choose the first ultrasound image as the reference image. The next step is to identify all other ultrasound

images from the hand being in resting position. The hand remains in this positions between two finger movements about one and a half seconds. As it follows, about 35 frames are gathered during this time (the system gathers valid ultrasound images at 20 Hz). Among these 35 images, one is chosen as the new compensation image. Since there is a short delay between the stimulus and the actual movement of the hand due to response time, the new compensation image has to be chosen carefully. In practice, the last of the 35 preselected images is chosen. This is the moment where the next movement of the stimulus model is just about to begin.

One problem with the probe motion compensation is, that due the shift of the feature points, some of them would move out of scope during the session. In that case, these feature points transfer wrong information and are therefore not considered when estimating the matrix K . Hence the algorithm has to be applied twice; the first time to evaluate all the features that transfer correct information through out the whole session, and the second time to extract the feature values from these feature points.

This also means that when applying the probe motion compensation, some information will get lost. In other words, the marginal feature points are not taken into account when evaluating the regression function.

Compensation On-Line

In the on-line version of the setup, however, the probe motion compensation is also used to monitor the movements of the subject's forearm. Similar to the off-line version, the compensation technique can only be applied with the hand being in the resting position. To determine the instant of time this is the case, the system either monitors the sequence of the virtual hand model or waits for manual input (for example, pressing a key on the keyboard). The optical flow algorithm is applied and the feature points are shifted.

The important step is to check if all feature points are still inside the field of view. If this is the case, the system shifts the feature points and continues to gather data, otherwise the systems stops and waits until all feature points are in the field of view again. In order to do so, the subject is asked to move the arm (that may unconsciously have been moved) back into the original position. As assistance, the offset values from the feature points are displayed on the screen. To be more precise, the average value in the horizontal direction and vertical direction are computed. This leads to two variables, ranging from about $-imagesize$ to $+imagesize$ (values in pixel). If these two variables are close to zero, no or only slight movement has been detected and the system does not stop. The higher the values, the higher the movement between to images.

5.4.2 Signal Filtering and Data Processing

In case of the on-line model, the output finger position values computed by the regression function are directly displayed by the virtual hand model. In order to avoid flickering and unintended finger movements of the hand model, a simple filter algorithm is applied. This filter calculates the mean value of the last three samples for each degree of motion. The hereby introduced delay can be neglected.

When analysing the data gathered in the preliminary experiments, the image features are

low-pass filtered using a Butterworth 5th-order filter.

In order to be independent from an installed version of Matlab, the required steps to estimate the regression matrix have also been implemented into the on-line version using a open source library called *EIGEN*. This is feasible, since the estimation of the regression function is basically just building the inverse of the design matrix. However, a disadvantage of this implementation is higher computation time.

6 Experiments Description

Despite some early tests, the setup presented in the previous section was basically used to perform two major series of experiments. Preliminary experiments were conducted on six healthy subjects to prove the general feasibility of the approach, and in the final experiments the system was tested on an amputee.

In all the experiments the subjects are instructed to perform a variety of single or multiple finger movements. The virtual hand model in figure 6.1 illustrates the twelve different hand motions including the resting position. Table 6.1 gives a detailed description.

Number	Movement
1	Resting position
2	Flexion of the pinkie
3	Flexion of the ring finger
4	Flexion of the middle finger
5	Flexion of the index finger
6	Rotation of the thumb
7	Adduction of the thumb
8	Flexion of pinkie and ring finger (gun shape 1)
9	Flexion of pinkie, ring and middle finger (gun shape 2)
10	Flexion of pinkie, ring, middle and index finger
11	Flexion of pinkie, ring, middle and index finger, adduction of thumb (lateral grasp)
12	Flexion of middle and index finger

Table 6.1: Different finger movements

As one can see, the twelve hand motions consist of single finger flexions or a combination of them. Only six degrees of motion are used in all movements, namely flexion of the pinkie, ring, middle and index finger as well as rotation and adduction of the thumb (movements 2,3,4,5,6 and 7 in figure 6.1).

6.1 Preliminary Experiments

The goal of the preliminary experiments is to make sure that the way of data gathering works on different subjects. Therefore the experiments were performed on six healthy subjects (5 male and 1 female between 22 and 39 years of age; 5 right-handed and one left-handed).

The subjects are asked to wear the cyberglove and then lie the arm relaxed on an orthopaedic support. A bench vice is used to fix the ultrasound transducer onto the forearm (transducer

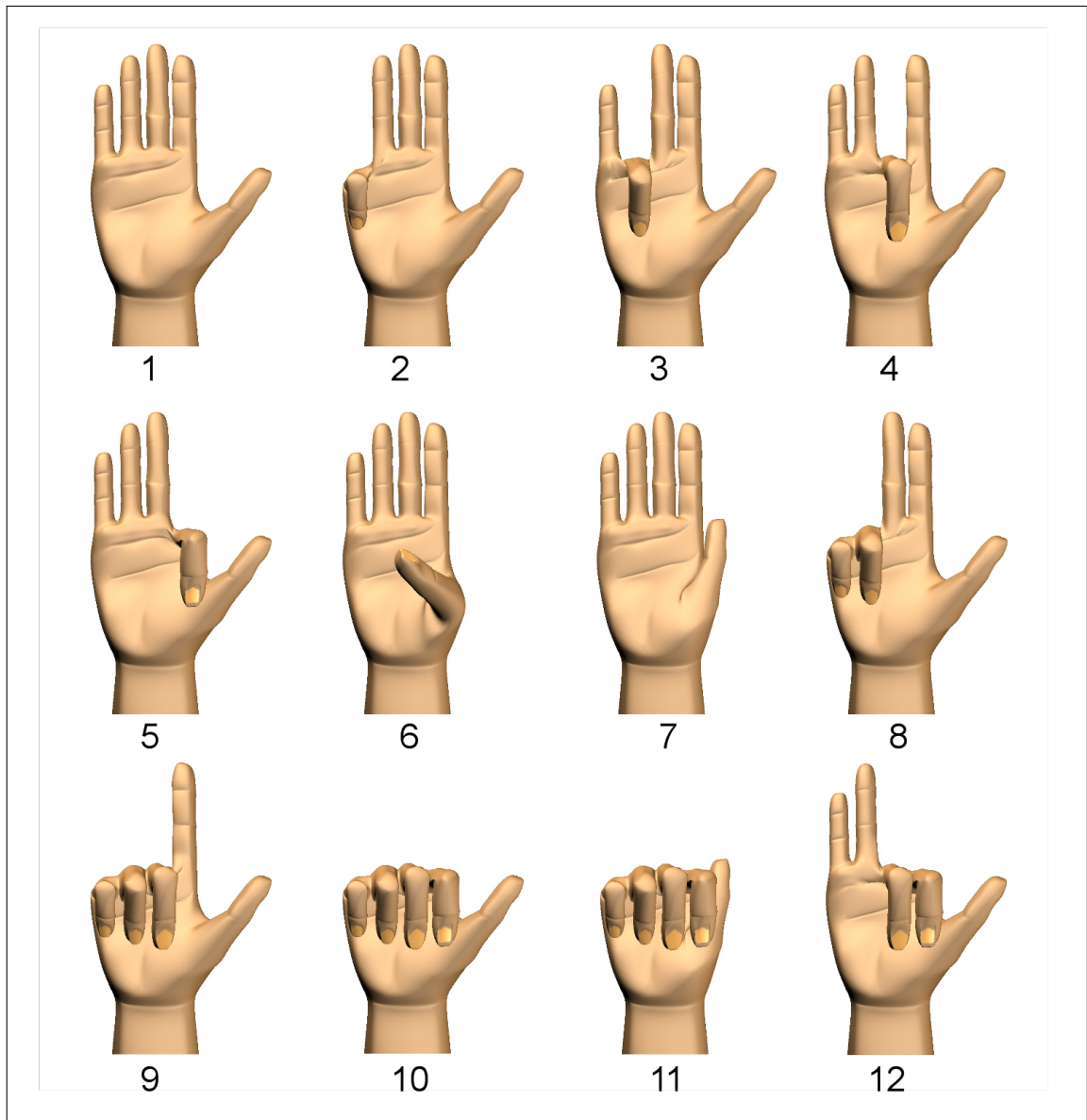


Figure 6.1: Different finger movements including resting position

position 1, near the wrist). Care has to be taken that the transducer is placed tight enough to maintain full contact with the skin. Standard ultrasound gel is applied to ensure the correct functionality of the ultrasound system. Figure 6.2 pictures the setup used in the preliminary experiments. One can see ultrasound system and how the transducer is fixed onto the forearm as well as the virtual hand model on the screen.

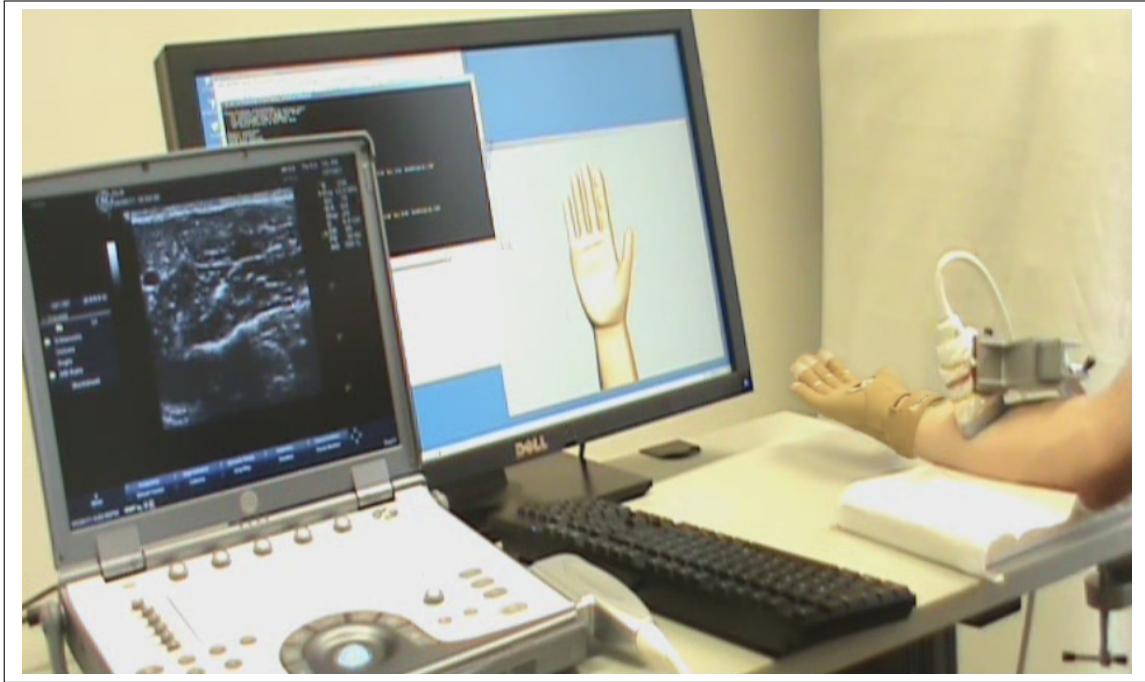


Figure 6.2: Setup used in preliminary experiments

The preliminary experiment is structured as follows: the subject is instructed to mimic different finger movements demonstrated by the virtual hand model. The set of finger movements consists of 10 different motions and the resting position (all movements displayed in figure 6.1 except motion 11). The single finger movements (2,3,4,5,6 and 7) are each displayed two times at three different speeds (2, 4 and 6 seconds). Similar to this, the multi finger movements (8,9,10 and 12) are shown three times at the same three speeds. This results in a total amount of 72 finger movements, which is called a *session*. Between the several movements, the hand remains in the resting position for about 1.5 seconds (also called period of rest). The 72 movements are shuffled beforehand so the subject is not aware of the sequence. With an average length of 4 seconds per movement, one session lasts about 7 minutes. Each subject does two sessions with a break of 10 minutes in between.

Before a session begins, the cyberglove has to be calibrated so each of the six glove sensors ranges between 0 and 1. In one session, approximately ultrasound are gathered. The results of the preliminary sessions are presented in section 7.1.

6.2 Final Experiments

In the final experiments the system was tested on a right handed amputee to substantiate the results from the preliminary experiments. Since an amputee is obviously not be able to wear a glove on the same arm the transducer is placed on, two variations of the on-line setup were applied. Therefore the subject either wears a left hand-handed glove while the transducer is placed on the stump, or the virtual hand's position values are used as ground truth. In the latter case, the subject has to mimic exactly the movements the hand model demonstrates with his imaginary limb. When using the left-handed glove, the subject perform synchronous movements with both the healthy and the imaginary hand.

Before after conducting the experiment, the subject is given an informative questionnaires, aimed to collect information about the subject's medical condition, level of phantom sensations, and usage of prostheses.

6.2.1 Experiment 1

The first experiment, in which no no glove is used, is divided into four phases:

- Phase 1: The subject is seated in front of the computer screen, the stump lying on a table. A virtual hand model, which is called the *teacher*, is shown on the screen, performing ten different finger movements (movements 2-11 in figure 6.1). The subject is asked to move the imaginary limb coherently, mimicking the finger motions on the screen. This phase is supposed to get the subject familiar with the system.
- Phase 2: Ultrasound gel is applied and the ultrasound transducer is placed onto the stump. The teacher hand model will again demonstrate finger movements and the subject is asked to mimic them again with the imaginary limb just, like in Phase 1. In this phase, the ultrasound system is gathering images.
- Phase 3: After the system generated the regression function, a second virtual hand model, called the *pupil* is displayed on the screen. The pupil hand model is used to illustrate the system's output finger positions on line. The subject's task to try and move the pupil hand model just like the teacher is moving.
- Phase 4: The teacher is turned off and the subjects is allowed to perform free movements and test the system as wanted.

In this case, the finger movements the teacher demonstrates are not shuffled beforehand. This should help the subject to be as synchronous with the teacher as possible.

6.2.2 Experiment 2

The second experiment, in which the subject wears a left-handed glove, is divided into three phases:

- Phase 1: The subject wears a left-handed glove on the intact hand. The teacher appears again and displays, after calibration, a replication of the intact hand. In this phase the subject should familiarise with the device and move the teacher into different positions with different velocities.

- Phase 2: Just like in the previous experiment, the transducer will again be placed on the stump. This time the subject is required to move both, the intact hand and the imaginary limb, synchronous. Similar to experiment 1, the ultrasound system is gathering images. In this section, the subject is allowed to decide which particular movements the system should be trained on.
- Phase 3: After the system has been trained, the pupil hand model appears as well and displays the intended movement of the imaginary limb. In this section, the subject is able to move freely. Phases 2 and 3 can be repeated as many times as wanted. This means that motions can be retrained if necessary. It is also possible to add new motions without interfering with already trained ones. Again, at the end of the experiment.

7 Results and Discussions

In this section, the results of the preliminary and final experiments described in section 6 are presented.

7.1 Preliminary Results

The aim of the preliminary experiments is to prove that the approach is working for different subjects. In order to substantiate the allegation that all human beings differ from each other, ultrasound images of the six subjects are shown in figure 7.1. The images in one row belong to the same subject but are recorded in different sessions. Since it is not able to place the ultrasound transducer exactly on the same spot in two sessions, small differences in the ultrasound images among the same subject are unpreventable. When comparing the subjects, one can clearly see differences in the images. However, some structures such as the pronator quadratis muscle or the radius bone (lower right area) occur in all images.

With each session lasting about 7 minutes, approximately 8000 samples per subject and session are gathered. In order to have an idea of the generality of this procedure, first the data set is randomly permuted for all subjects. The training set consists of a certain subset. After the matrix K is estimated, the system is tested for prediction on the test set. The prediction is repeated for 100 times, each time at different permutation. For each subject, the mean absolute error (ERR), normalised square-root mean-square-error (NRMSE) and correlation coefficient (CC) are displayed, obtained by linear regression on each finger position.

The best results are obtained when the training set consists of half of the full dataset and as one would expect, the error increases as the training set is reduced. Table 7.1 summaries the results obtained when the training set consists of a twelfth of the full dataset. In this case K is estimated over approximately $8000/12 = 666$ samples. This suggests that as few as $666/20 = 33$ seconds of training might be enough (note that the system gathers data at about 20 Hz). One also can see that the thumb adduction and thumb rotation uniformly remain the hardest motions to predict.

Table 7.2 lists the results for the six subjects after applying the probe motion compensation. The same data set with identical permutations and number of predictions is used. Again, errors and correlation coefficient are reported. As one can see, regression still holds and the results can be improved over all subjects.

On average, comparing the two tables 7.1 and 7.2, ERR decreases about 0.093 ± 0.023 and NRMSE about 0.12 ± 0.03 . The average correlation coefficient increases about 0.23 ± 0.045 . Figures 7.2, 7.3 and 7.4 display these results. The errorbars denote the mean and standard deviation for the errors and correlation coefficient obtained on the 100 permutations.

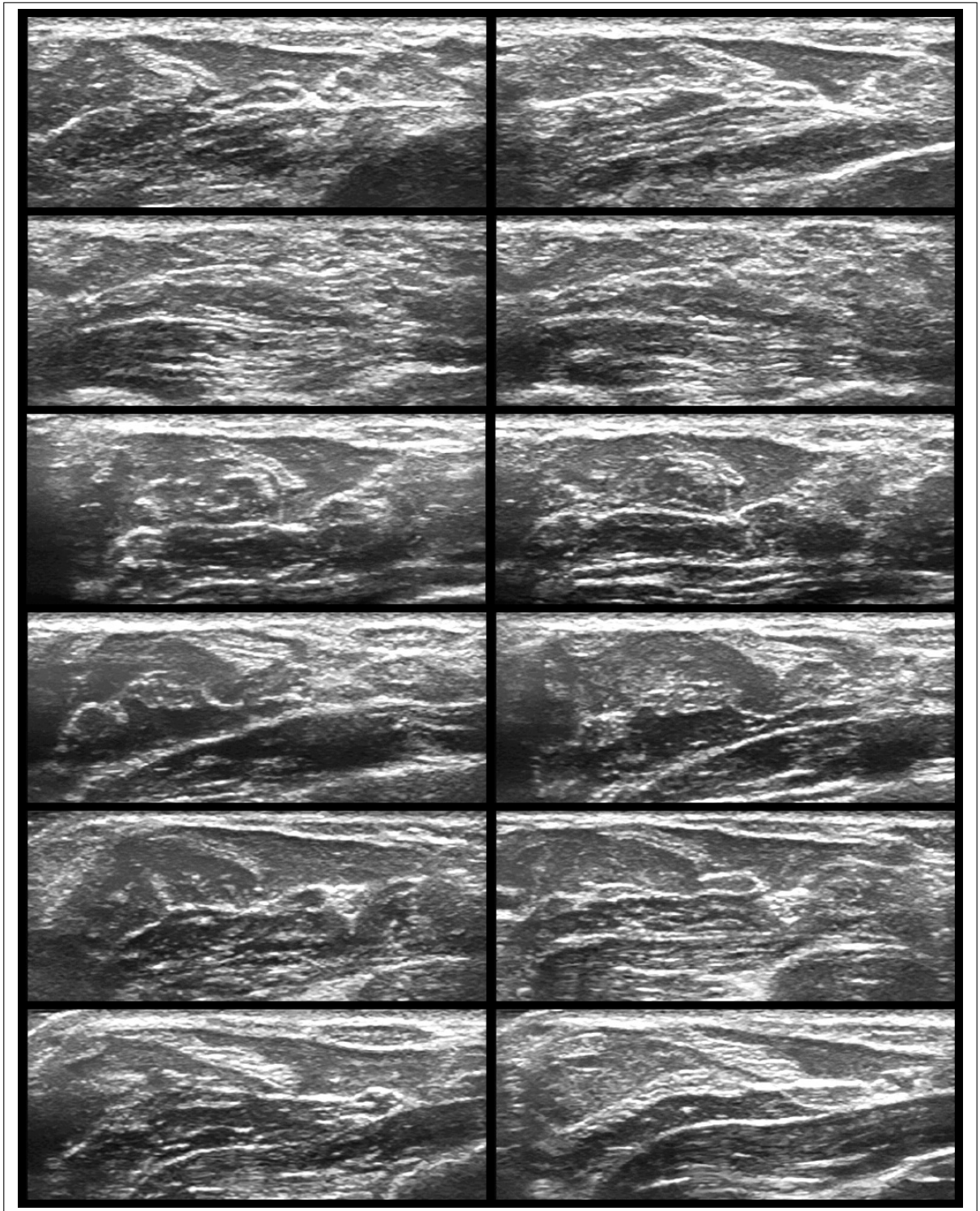


Figure 7.1: Comparison of ultrasound images of different subjects

subject		pinkie	ring	middle	index	th. rot.	th. add.
1	ERR	0.1617	0.1692	0.1277	0.1258	0.2349	0.1973
1	NRMSE	0.2080	0.2516	0.1858	0.1827	0.1797	0.2144
1	CC	0.7612	0.7442	0.8144	0.7868	0.5012	0.4163
2	ERR	0.2751	0.1972	0.1064	0.1324	0.1553	0.2383
2	NRMSE	0.3004	0.2336	0.1430	0.1665	0.1893	0.2852
2	CC	0.4615	0.5647	0.7407	0.7101	0.5718	0.4096
3	ERR	0.2174	0.1343	0.1077	0.1074	0.1971	0.1378
3	NRMSE	0.2211	0.1883	0.1302	0.1382	0.1284	0.1607
3	CC	0.5919	0.6846	0.7768	0.7520	0.6345	0.3839
4	ERR	0.1598	0.1027	0.0950	0.1120	0.1907	0.2518
4	NRMSE	0.2305	0.1863	0.1727	0.1644	0.2602	0.2801
4	CC	0.6838	0.7902	0.7988	0.7848	0.4241	0.3742
5	ERR	0.2187	0.2074	0.1529	0.1316	0.1637	0.1575
5	NRMSE	0.3195	0.3347	0.2780	0.2370	0.1718	0.1976
5	CC	0.6603	0.6332	0.6728	0.7157	0.5683	0.4677
6	ERR	0.1643	0.1615	0.1354	0.1524	0.1638	0.1867
6	NRMSE	0.2154	0.2294	0.1788	0.1993	0.1672	0.2202
6	CC	0.7197	0.7170	0.8044	0.7627	0.5824	0.4416

Table 7.1: Error results obtained by linear regression on each degree of motion: mean absolute error (ERR), normalised square-root mean-square-error (NRMSE) and correlation coefficient (CC).

The improvement after using the probe motion compensation can also be seen in figure 7.5. In this case, the system was trained on the first half of the samples and tested on the second half. The figure shows a snippet of the prediction results, where the effect of the probe motion compensation is clearly visible. The black line represents the target values for one degree of motion, namely the flexion of the middle finger. The corresponding predicted values with probe motion compensation (blue dotted line) and without (red dotted line) are displayed as well. As one can see, after some time the red line starts to drift away from the target line, whereas the blue line better correlates.

subject		pinkie	ring	middle	index	th. rot.	th. add.
1	ERR	0.0707	0.0717	0.0548	0.0539	0.1021	0.0848
1	NRMSE	0.0894	0.1045	0.0781	0.0767	0.0767	0.0890
1	CC	0.9387	0.9366	0.9574	0.9492	0.7977	0.7316
2	ERR	0.1156	0.0851	0.0512	0.0592	0.0677	0.1030
2	NRMSE	0.1234	0.0990	0.0680	0.0733	0.0802	0.1201
2	CC	0.7777	0.8442	0.9161	0.9133	0.8496	0.7126
3	ERR	0.1102	0.0655	0.0547	0.0548	0.0922	0.0655
3	NRMSE	0.1098	0.0913	0.0653	0.0696	0.0594	0.0758
3	CC	0.8200	0.8868	0.9261	0.9141	0.8730	0.6471
4	ERR	0.0688	0.0460	0.0395	0.0481	0.0906	0.1121
4	NRMSE	0.0975	0.0820	0.0705	0.0690	0.1205	0.1209
4	CC	0.9100	0.9463	0.9559	0.9489	0.6938	0.6536
5	ERR	0.0948	0.0884	0.0666	0.0625	0.0700	0.0667
5	NRMSE	0.1364	0.1405	0.1191	0.1104	0.0726	0.0825
5	CC	0.8986	0.8880	0.9028	0.9097	0.8526	0.7771
6	ERR	0.0659	0.0661	0.0538	0.0569	0.0618	0.0759
6	NRMSE	0.0833	0.0910	0.0686	0.0714	0.0608	0.0858
6	CC	0.9374	0.9330	0.9616	0.9567	0.8895	0.7778

Table 7.2: Error results obtained after applying probe motion compensation: mean absolute error (ERR), normalised square-root mean-square-error (NRMSE) and correlation coefficient (CC).

As one would expect, the regression accuracy highly depends on the number of training samples. To make this clear, the system is trained on different subsets (depending on the step size n) of the initial, non-shuffled data. As usual, the system is tested on the remaining samples. The step sizes for the training set are 2, 4, 6, 8, 10 and 12 (for example, with step size $n=6$, the system is trained on every sixth sample). Table 7.3 reports the errors and correlation coefficient. It is obvious that the errors increase as the training size decreases. This effect holds uniformly for all subjects. Figure 7.6 illustrates the regression accuracy for a system with training set step sizes 2, 6 and 10. The figure shows the target and predicted values for one degree of motion, namely flexion of the pinkie.

step size		pinkie	ring	middle	index	th. rot.	th. add.
n = 2	ERR	0.0278	0.0270	0.0227	0.0246	0.0262	0.0285
	NRMSE	0.0323	0.0342	0.0267	0.0289	0.0236	0.0295
	CC	0.9901	0.9901	0.9940	0.9927	0.9821	0.9679
n = 4	ERR	0.0300	0.0289	0.0247	0.0267	0.0282	0.0302
	NRMSE	0.0348	0.0366	0.0290	0.0315	0.0255	0.0315
	CC	0.9886	0.9886	0.9929	0.9913	0.9791	0.9635
n = 6	ERR	0.0337	0.0333	0.0278	0.0297	0.0323	0.0353
	NRMSE	0.0406	0.0432	0.0332	0.0352	0.0293	0.0366
	CC	0.9845	0.9842	0.9908	0.9892	0.9727	0.9515
n = 8	ERR	0.0436	0.0423	0.0350	0.0373	0.0389	0.0450
	NRMSE	0.0518	0.0538	0.0429	0.0450	0.0362	0.0483
	CC	0.9752	0.9757	0.9846	0.9822	0.9587	0.9190
n = 10	ERR	0.0617	0.0607	0.0475	0.0525	0.0568	0.0650
	NRMSE	0.0755	0.0811	0.0588	0.0649	0.0538	0.0714
	CC	0.9488	0.9467	0.9716	0.9643	0.9153	0.8461
n = 12	ERR	0.2065	0.1341	0.1729	0.1202	0.1286	0.1712
	NRMSE	0.2547	0.1776	0.2152	0.1491	0.1218	0.1895
	CC	0.6510	0.8012	0.7432	0.8536	0.7051	0.4929

Table 7.3: ERR, NRMSE and CC for a system with various training set size (n denotes the step size)

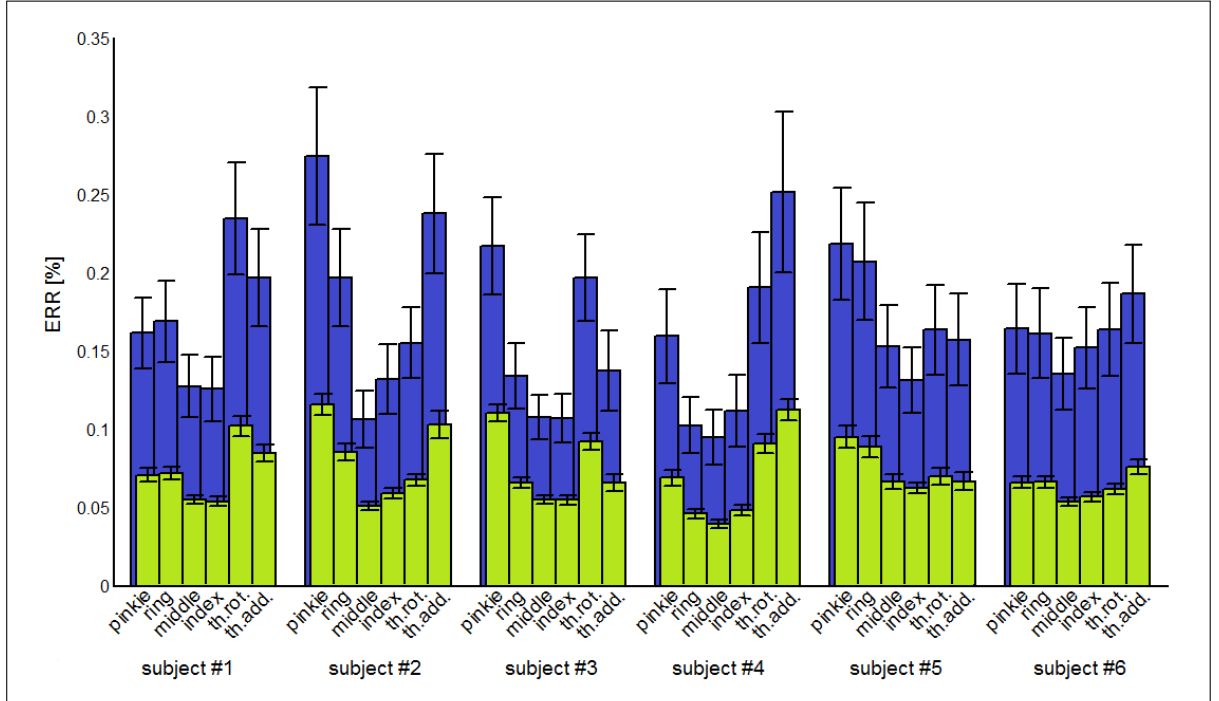


Figure 7.2: Comparison of ERR before (blue bars at back) and after (green bars at front) probe motion compensation. Bars denote the mean and standard deviation obtained on the 100 permutations.

7.1.1 Correlation Coefficient

In order to illustrate which parts of the ultrasound image show high correlation with certain degrees of motion, a data set, large enough to compute the correlation coefficient for a very dense grid of feature points, was gathered. Figure 7.7 shows the correlation coefficient for each feature point (for the purpose of illustration, the correlation coefficients of the three features α , β and γ have been averaged out for each feature point). As one can see, several specific areas of high correlation for the first four degrees of motion exist. In contrast to this, for the movements of the thumb, only few feature points show very high correlation. In this case, when comparing the correlation coefficient map with the ultrasound image (gathered on the mid-arm position of the forearm), the area with highest correlation represents one of the extensor muscles, namely abductor pollicis longus (see also figures 4.8 and 4.9).

7.2 Final Results

Before the results of the two experiments are discussed, the findings of the pre-experiment questionnaires are presented.

7.2.1 Questionnaires

In order to have an idea of the subject's medical condition, some basic information is listed in table 7.4.

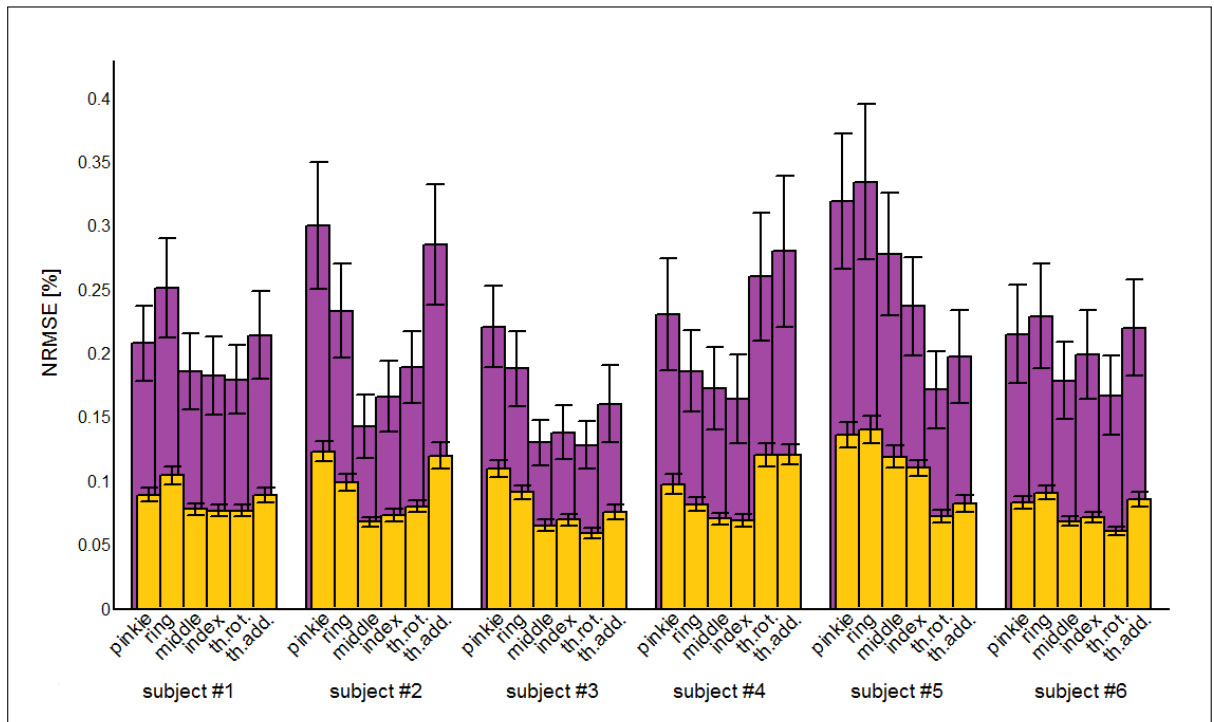


Figure 7.3: Comparison of NRMSE before (purple bars at back) and after (orange bars at front) probe motion compensation. Bars denote the mean and standard deviation obtained on the 100 permutations.

Date of Birth	24.09.1967
Gender	Male
Age	43
Weight	73
Height	183 cm
Occupation	Employee
Hobbies	Sports
Sports	Swimming, Biking, Running

Table 7.4: Pre-experiment questionnaire

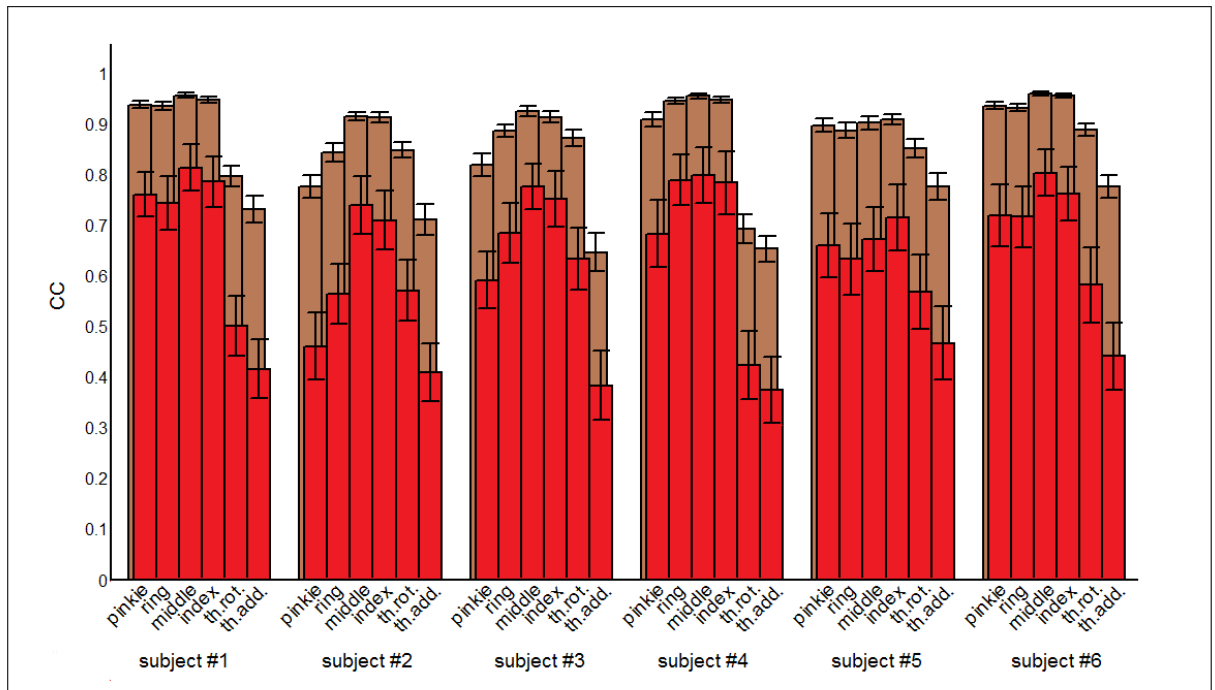


Figure 7.4: Comparison of CC before (red bars in front) and after (brown bars at back) probe motion compensation. Bars denote the mean and standard deviation obtained on the 100 permutations.

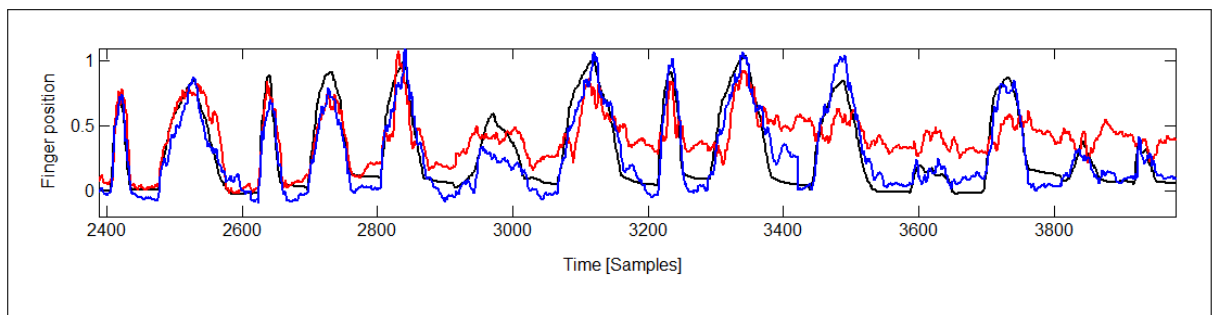


Figure 7.5: Target values (black line) and predicted values with probe motion compensation (blue dotted line) and without (red dotted line)

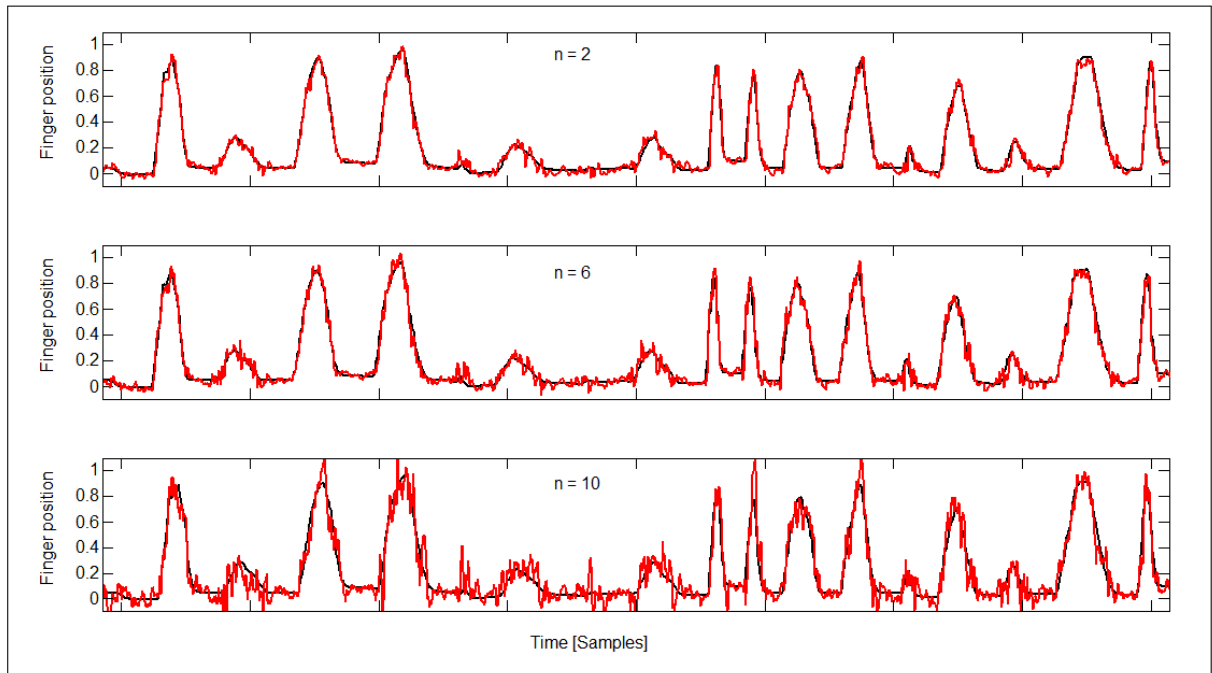


Figure 7.6: Regression accuracy depends on training set size. The system was trained on every second (first row) every sixth (second row) or every tenth (third row) sample of the data set

The patient participating in this experiment lost his hand due to an operation. Some details concerning these circumstances are shown in table 7.5.

operation details	
Date	April 2008
Type	Transradial distal
Cause	Hand tumor
Side	Right
Length of remaining limb	About 30 cm
Dominant hand before operation	Right
Dominant hand now	Left

Table 7.5: Pre-experiment questionnaire

Before the experiment started, the muscle activity in the remaining limb was examined by both palpation and monitoring the ultrasound images. In both cases a clear activation signal from the subjects muscles for all six degrees of motion could be observed.

Furthermore the subject describes the phantom limb sensation as rigid, wet, or like in a tight glove or concrete block. The subject relates them to stress, witnessed violence, physical injury, or cold and the sensations last for a few hours per day. The pain can be alleviated with hot baths and eating bananas (especially the latter one by exploiting the placebo effect). The subject also

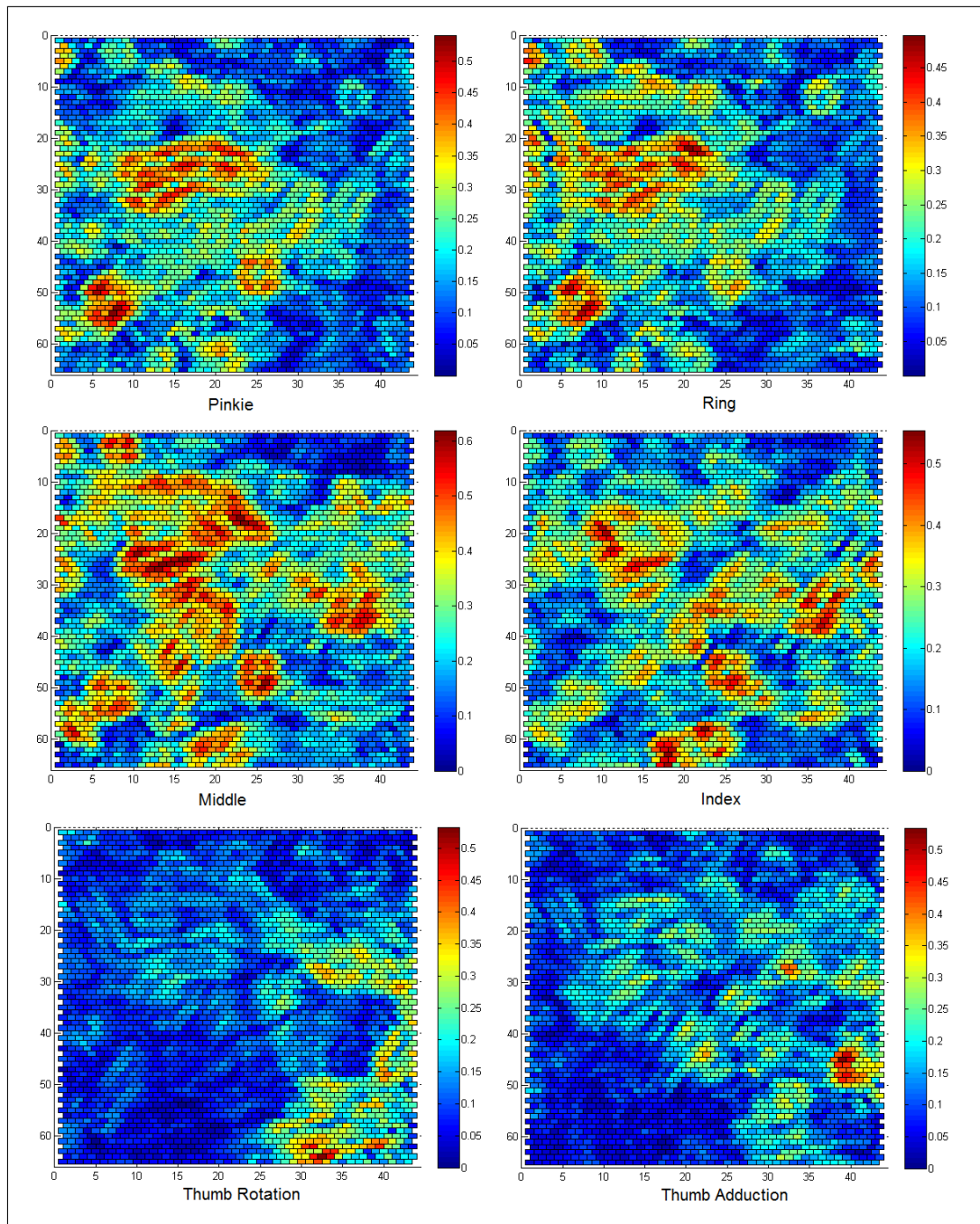


Figure 7.7: Correlations between features and degrees of freedom

Type of prosthesis	Usage of prosthesis
Cosmetic	Permanently
Body-powered	Permanently
Myoelectric, open/close	Sometimes
Myoelectric, polyarticular	—

Type of prosthesis	Average usage of prosthesis at work
Cosmetic	1-2 days per week
Body-powered	4-6 days per week
Myoelectric, open/close	Few hours per month
Myoelectric, polyarticular	—

Type of prosthesis	Average usage of prosthesis during free time
Cosmetic	1 evening per week
Body-powered	1 afternoon per week
Myoelectric, open/close	—
Myoelectric, polyarticular	—

Type of prosthesis	Advantages
Cosmetic	Weight, look
Body-powered	Emotional
Myoelectric	Great force

Type of prosthesis	Disadvantages
Cosmetic	Skin issues, repairs
Body-powered	Skin issues, repairs
Myoelectric	Weight, bad reliability

Table 7.6: Usage of prosthesis

pointed out that the brain has learned to suppress any feelings, including pain, coming from the residual limb.

7.2.2 Experiment 1

In phase 1 of the first experiment the subject made itself familiar with the setup. In phase 2 however, difficulties emerged when fixating the ultrasound transducer to the stump. Due to the fact the intended positioning of the forearm (like in the preliminary experiments; see figure 6.2) was not comfortable for the subject, the fixation of the transducer had to be changed. The Image 7.8 shows the new and rather improvised fixation of the arm. Unfortunately, using this new and more comfortable position, it was not able to keep the transducer still on the stump and therefore the change of the setup had great impact on all results throughout the experiments.

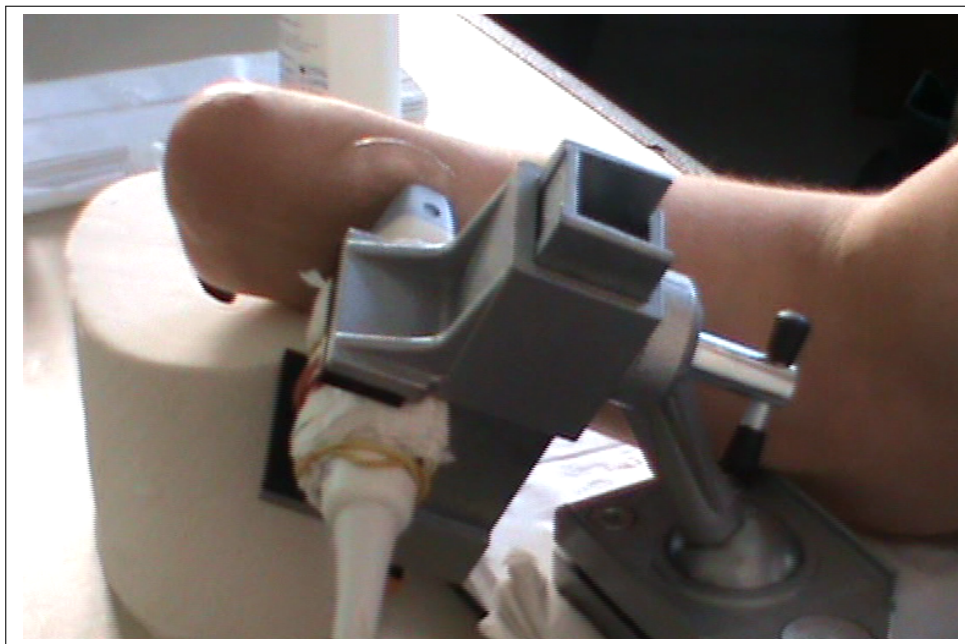


Figure 7.8: Fixation of the forearm

Other changes concerning the teacher hand model had to be made. The experiment originally started with 10 movements, each movement lasting 5 seconds. The duration of each movement then was reset to 2 seconds because the subject did not feel able to move his imaginary limb that slowly. Also the amount of different movements was reset to six, namely the flexion of the four fingers and the two thumb movements (movements 2-7 in figure 6.1).

Despite the changes mentioned above, the subject was able to conduct the experiments as scheduled conducted several short sessions. Figure 7.10 pictures the subject in the data gathering phase. One can see how the teacher demonstrates the index finger flexion. The subject tries to mimic the movement with its imaginary limb.

In figure 7.10 the subject is testing the system. As can be seen, the pupil hand model on the right side, which is displaying the predicted finger positions on-line, is moving coherently with the

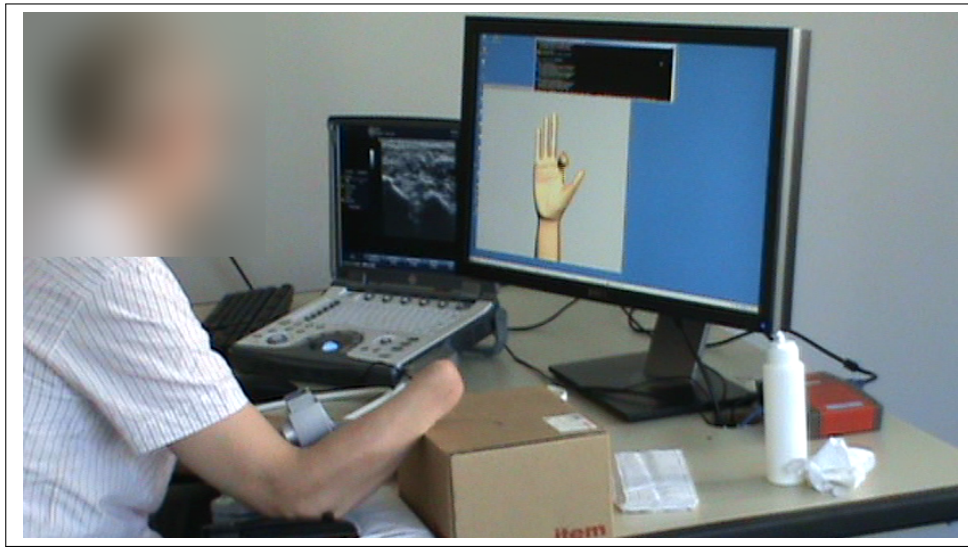


Figure 7.9: Experiment 1: The subject is training the system by mimicking the movements on the screen

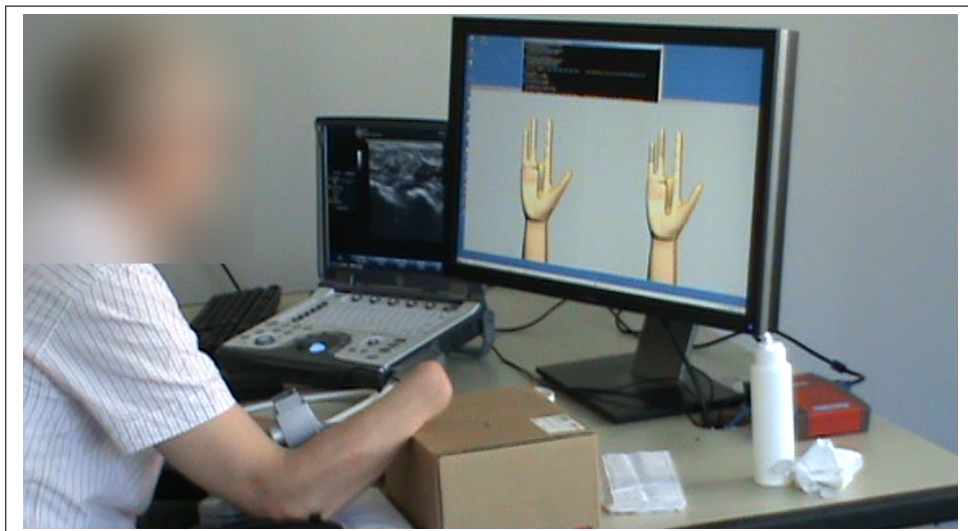


Figure 7.10: Experiment 1: Testing phase

teacher.

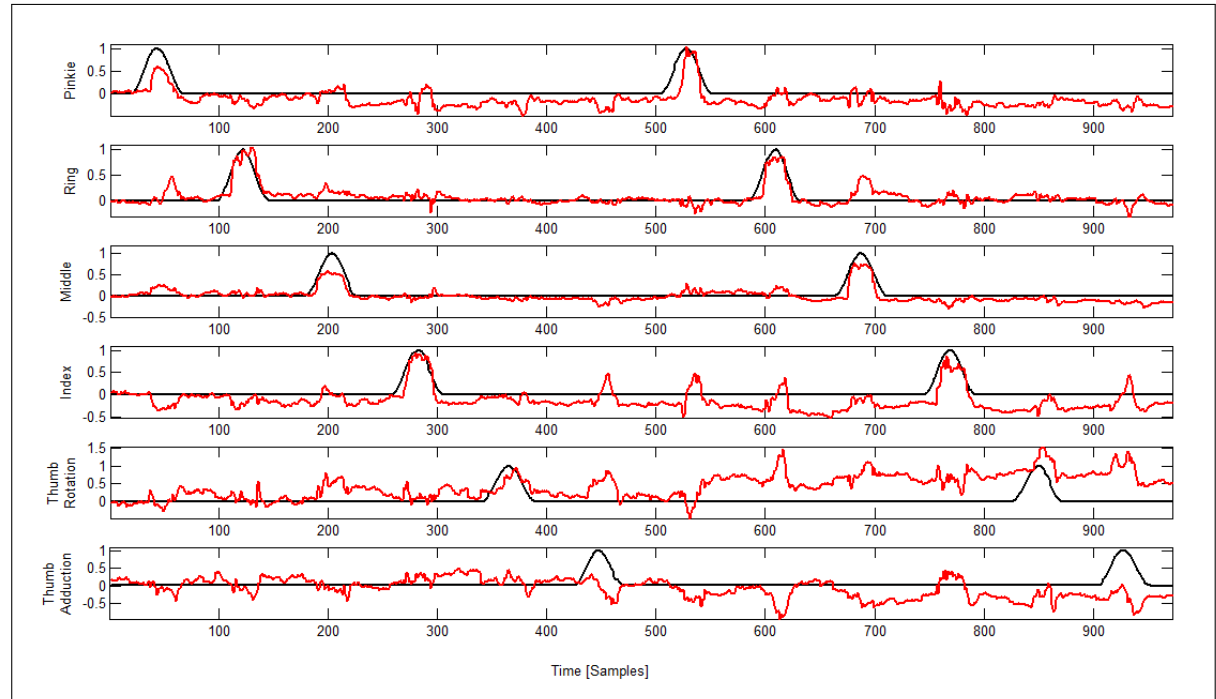


Figure 7.11: Results for all six degrees of motion. Target position values represented by the black line, predicted values by the red line

Figure 7.11 shows the best results obtained from the first experiment. Approximately 1000 frames are gathered. The black line represents the target values, the red dotted line the predicted values. As apparent, a clear correlation for the first four degrees of motion (pinkie, ring, middle, and index finger) can be attained, but the system is not able to correctly predict the thumb position.

7.2.3 Experiment 2

In the second experiment, the subject is allowed to decide which particular movements the system should train. Again, several short session from this experiment were performed. Figure 7.12 pictures the subject wearing the left-handed cyberglove. This time, the teacher replicates the movements of the left hand while the subject moves the imaginary limb coherently.

The testing phase of the second experiment is displayed in figure 7.13. The subject flexes the pinkie and system predicts the movement correctly.

Again, the best results are shown in figure 7.14, a session with approximately 500 samples. Apparently the subject only trained the flexion of the pinkie and index finger. Whereas the system is able to correctly predict the pinkie finger positions (see also figure 7.13) it more ore less fails on the index finger.

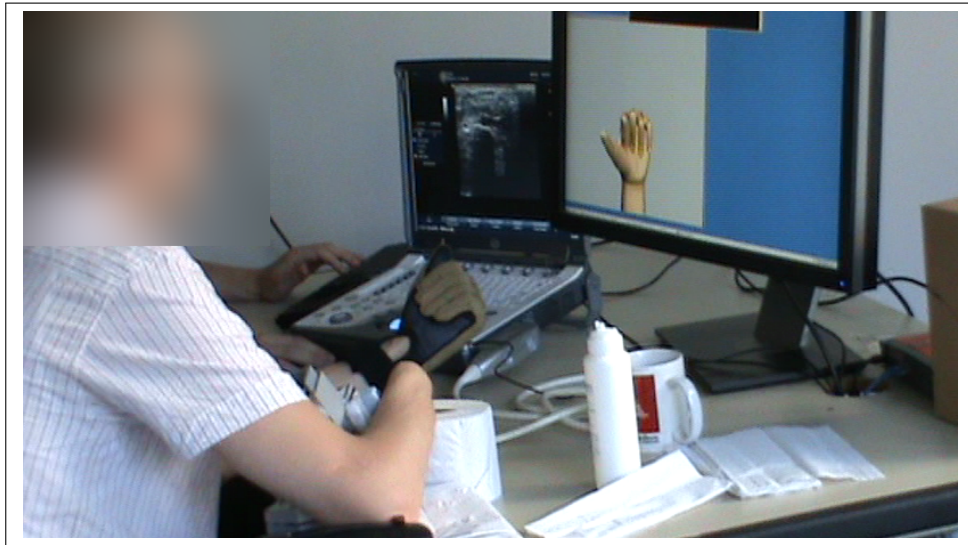


Figure 7.12: Experiment 2: The subject trains the system moving the imaginary limb coherently with the left hand

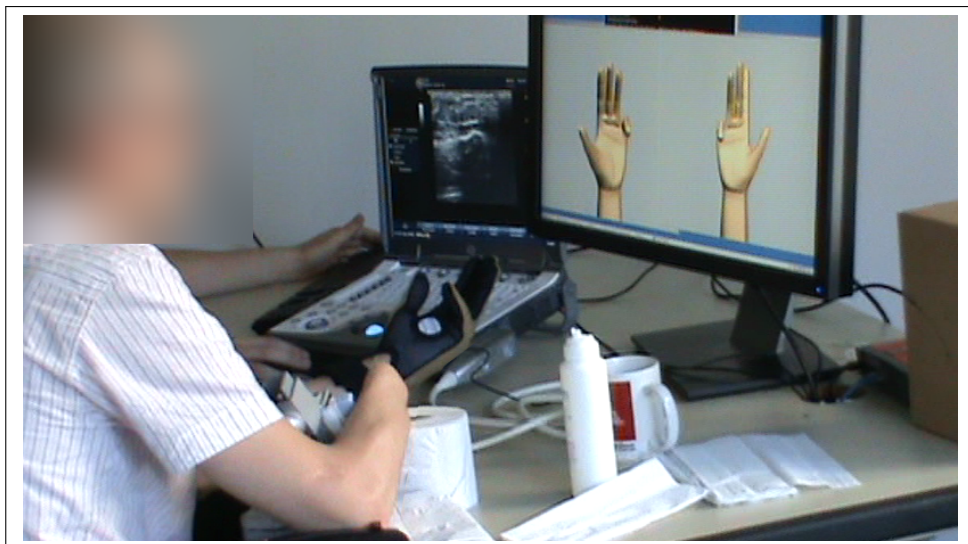


Figure 7.13: Testing phase of the second experiment

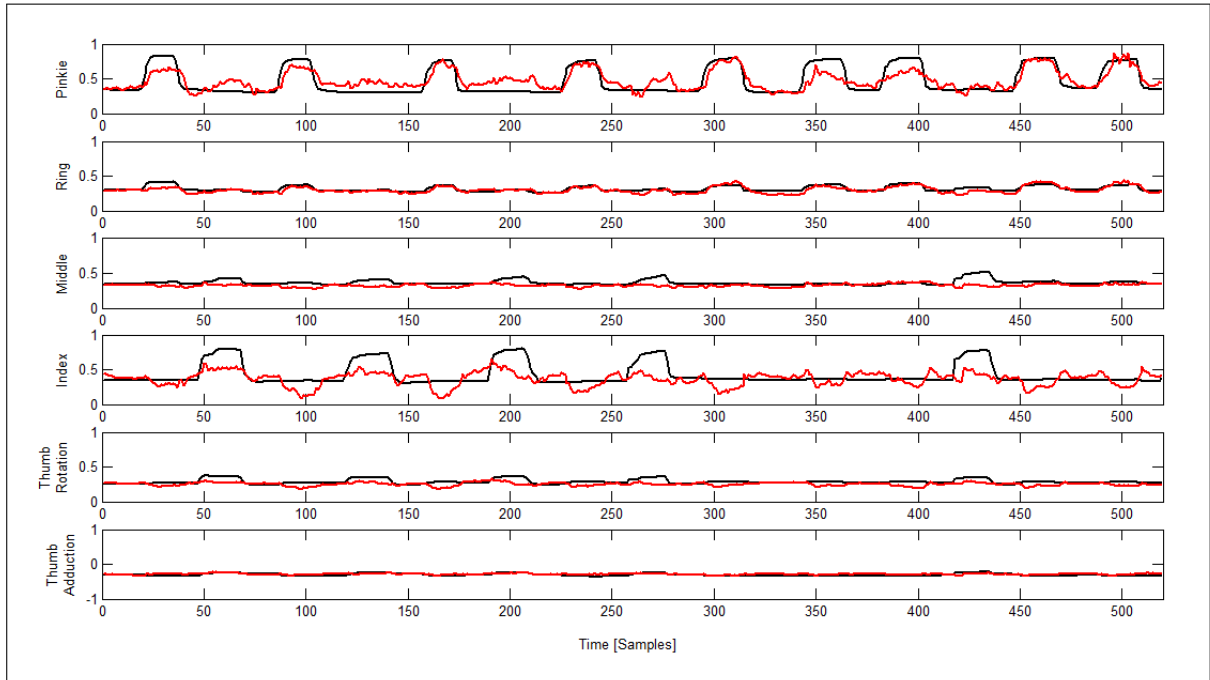


Figure 7.14: Results for all six degrees of motion. Target position values represented by the black line, predicted values by the red line

It should be noted again that most of the errors in the two experiments stem from the changed transducer fixation and it is believed that better results can be obtained after the setup has been improved.

7.3 Feature Properties

A detailed description of features extracted from all the images is given in sections 5.3.1 and 4.4.2. In this section, results from an estimation grid search to find the optimal setting of the feature point grid are presented. Two parameters, namely the radius of each feature area and the distance between grid points have an influence on the speed and quality of the performance.

The data from one session consisting of about 11800 frames have been taken to find the optimal adjustments. The data set was tested on high correlation coefficient between image features and finger movements and the mean absolute error. Therefore, as it is customary, the data set was divided in two parts, one for training and the other one for testing. Either the training set consists of the first half of the data set and the testing set on the other half, or the data set has been shuffled before and has then again been equally partitioned. The radius ranges from 5 to 70 pixels and the distance between grid points from 20 to 110 pixels, both with a step size of 5 pixel. This leads to a total number of $19 \times 14 = 266$ variations.

Figure 7.15 shows the map for the correlation coefficient between image feature points and finger movements. Here the system was trained on the first half and tested on the second

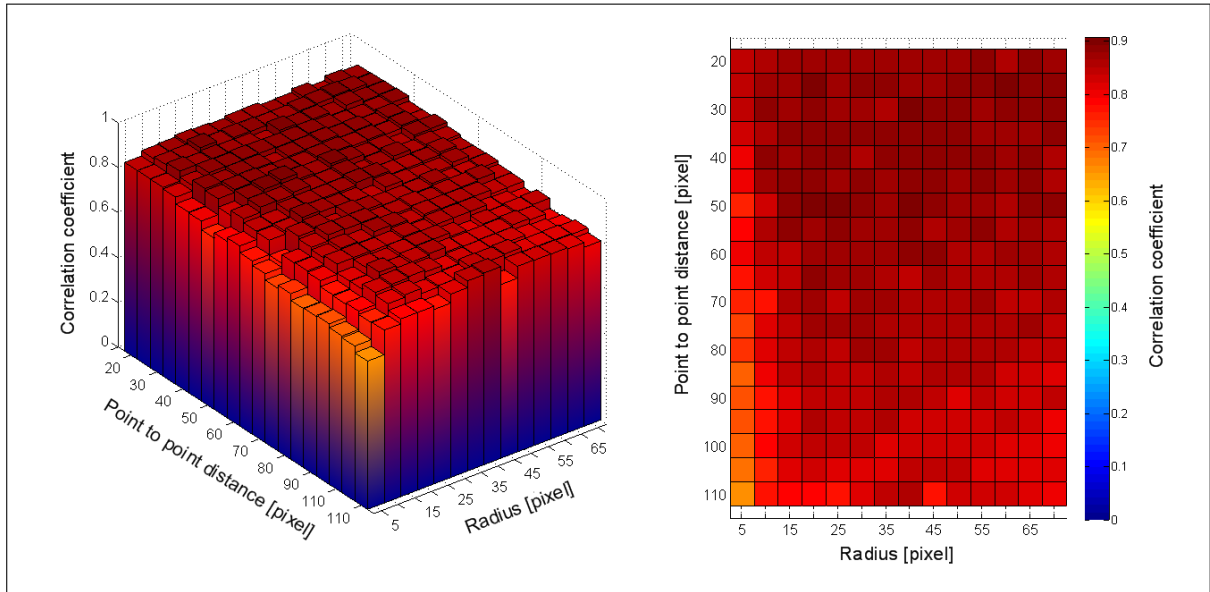


Figure 7.15: Correlation coefficient between image feature points and finger movements. The system was trained on the first half of the data set and tested on the other half

half of the data set. It can be seen that the more feature points, the higher the correlation coefficient. It is also evident that the radius does not have that much influence. The corresponding absolute mean error is displayed in figure 7.17. There is no clear convergence for the error visible but still the absolute mean error is smaller with a high number of feature of points.

The same tests have now been done with shuffling the data set before hand. Since shuffling does not guarantee that the samples are somehow split equally into training and test set (imagine the worst case where the index finger movement is not trained at all because no such samples are in the training set) the shuffling was repeated 100 times and the average of the results for both the correlation coefficient and the absolute mean error are taken (see figure 7.16 and figure 7.18).

It is obvious that this leads to a smooth distribution and a clear convergence for both parameters. Another side effect of shuffling the data set is that the correlation coefficient is much more higher in comparison to the system trained on the first half and tested on the second half. This is mainly due to the fact the similar samples were trained and tested and the systems generalises well. In this case, the movement of the ultrasound transducer does not have an impact on the performance.

7.4 Discussions

Since the estimation of the regression matrix K can be computationally costly for large training sets, the implementation of the least mean squares algorithm was considered. As described in section 4.3.2, the algorithm works as an adaptive filter and considers one sample at a time. However, implementing this algorithm did not show the desired effect of creating the estimation matrix on the fly. A more promising approach could be the use of regularised least squares (see section 4.3.2).

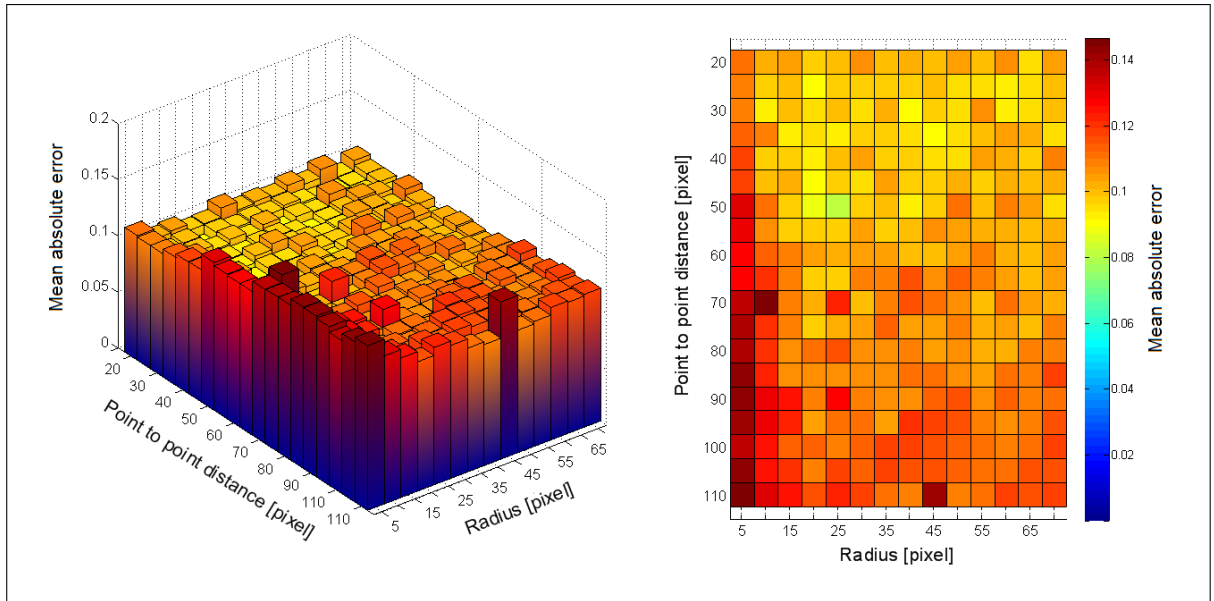


Figure 7.16: Mean absolute error for the system trained on the first half of the data set and tested on the other half

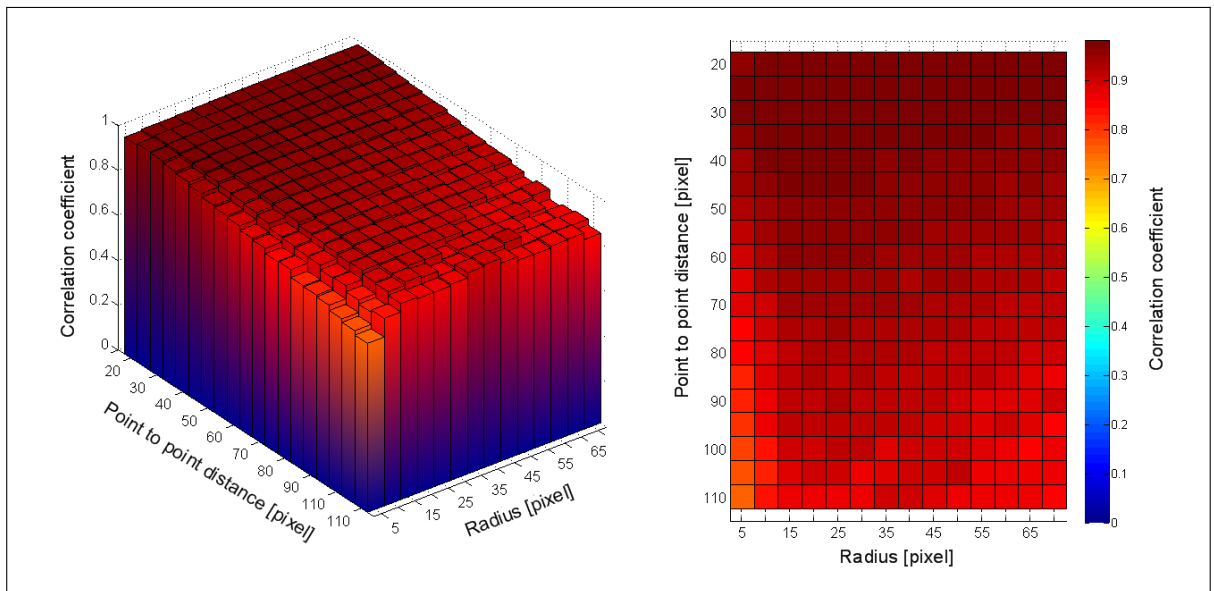


Figure 7.17: Correlation coefficient between image feature points and finger movements with shuffling the samples before training the system

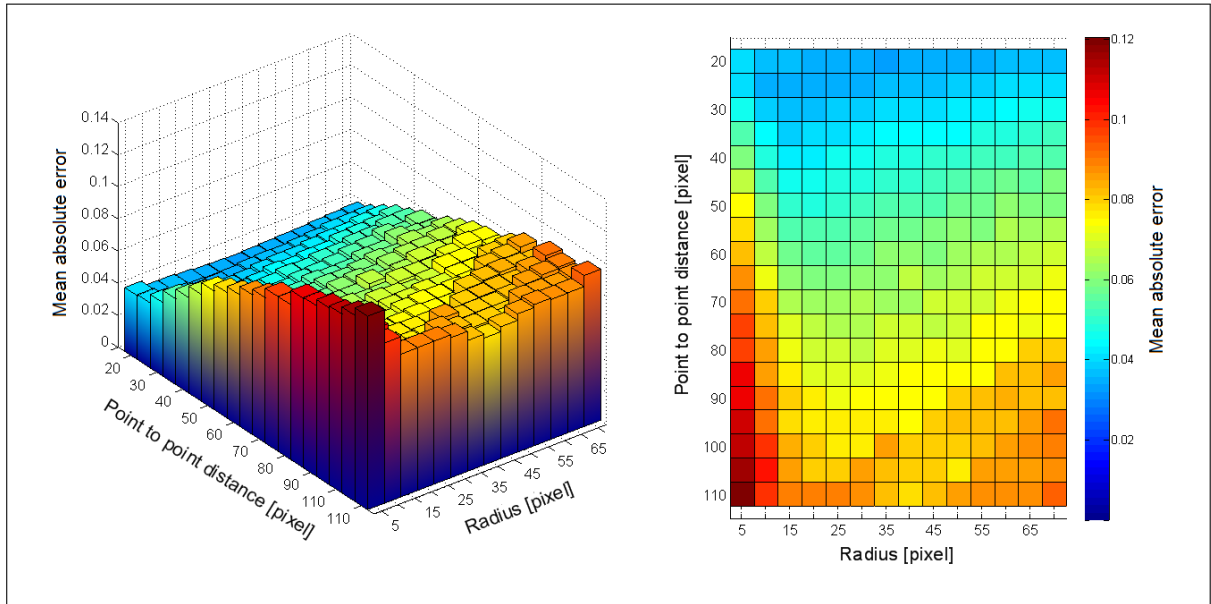


Figure 7.18: Mean absolute error for the system. Samples were shuffled before training

The possibility for amputees to use this system depends on several circumstances. Enough musculoskeletal potential is required for the system to work fine. The remaining muscle and tendon movements have to be distinguishable. For example, changes in the ultrasound image when flexing the pinkie must be different from changes when flexing the ring finger. This of course usually holds for healthy subjects but one can not assure this to be true for amputees. Therefore the optimal transducer position has to be estimated. The degree of musculoskeletal activity also depends on the date of the amputation and whether the subject uses myoelectric prostheses.

The only drawback with the on-line setup seems to be the optical flow calculation which can extend the computation time. But since the optical flow is required only between two finger movements, when a new reference frame is detected, the hereby induced delay can be neglected.

This system aims only at prediction the finger positions. When trying to predict as well force and/or movement velocity, a number of points have to be considered and this would imply much higher degree of complexity. Another question that arises is, when training the system, how many samples are needed at least, and, with a greater degree, how many samples are allowed to ensure the system to generalise well. In other words, how many different finger movements can be distinguished by the system at all. In all experiments, no limitations to the system whatsoever are reported. Moreover, this also highly depends on the dimensionality of the system (see also section 4.3.1).

8 Conclusion and Future Work

The experimental results show clearly that ultrasound imaging of the human forearm can be effectively used to reconstruct the position of the fingers in sub-degree precision. Thus the presented approach could be a novel control system for prosthetic hands and be a replacement of surface electromyography.

Probably the most interesting outcome of the experiments conducted in this work, is the linear relationship between image features and finger positions. This holds uniformly for both healthy subjects and, at least one, below-elbow amputee. The relationship is also robust with respect to relative motion between the ultrasound transducer and the subject's forearm. To compensate this inevitable movement, a simple optical-flow based technique is implemented.

The linear relationship allows the whole system to be implemented on-line. Therefore it is claimed that this approach builds the base for the control of dexterous hand prostheses. Furthermore the system is able to visualise the phantom limb of amputees and can be used as a tool for neuroscientists to treat phantom pain.

One advantage of the presented approach is that it shows complete disregard of the anatomical structures. In other words, the equally spaced grid stays uniformly for all subjects, and to a greater degree, to all transducer positions. Actually, in order to fully understand the phenomenon of the linear relationship, a model of the musculoskeletal structure, its motion and the view across a section of it would be required, but this is not the focus of this thesis.

One drawback of the approach so far is the inevitable movement of the transducer on the forearm skin. But not only the transducer is sliding on the skin, the elastic skin is moving with respect to the musculoskeletal structures below. In both cases, the introduced optical-flow based technique is able to compensate this shift to a certain extend. If the transducer moves too much, too many feature points move out of scope and the approach may be disrupted.

This problem could be largely solved by improving the setup, especially a better fitting and more convenient ultrasound probe. A more concave-shaped transducer would induce less movement. The transducer also might be steadily fixed on the forearm via a belt. A more sophisticated solution could be a glove or stocking with single transducers sewn atop, placed according to optimal muscular activity. Of course, a long-term goal could be the optimal solution: a miniaturised and wearable system able to control a polyarticular, dexterous hand prostheses.

Another interesting point is that of finding features which are less sensitive to small changes in the ultrasound image. So far, least mean squares method is used to estimate the regression function. But as presented in section 4.3.2, regularised least squares could be used as a more robust approach. Even though linear regression is a very fast and easy applicable approach, it has significant limitations in terms of pattern recognition. More sophisticated machine learning

methods, such as support vector machines or neural networks, could be implemented to break the linearity.

As the experiments in section 7.3 have revealed, the radius of the feature point area and, to a greater degree, the number of feature points, do have an effect on the performance of the system. In other words, there is a clear trade-off between the number of feature points and the system's regression accuracy. Once, a minimum set of feature points is found, local correlation with single finger movements could be evaluated. If it is possible to apply some sort of source separation, this could be used to enforce composition of these movements into multifinger motions.

Ultrasound imaging of the forearm can find different applications, both in virtual reality and real environments, including the control of mechanical and virtual hands or teleoperation. Ideally, the subject is wearing some sort of glove or stocking with the ultrasound sensors to obtain maximum mobility. Nevertheless, rehabilitation of upper-limb amputees is the most promising way ahead. Another advantage of the method is, that the location of the ultrasound transducer does not influence very much the experiment. Examining the ultrasound images reveals musculoskeletal activity along the whole forearm, including flexor and extensor muscles.

So far, only six degrees of motion are used to build the regression model. Extending this number by also considering the remaining sensors of the cyberglove, the number of different finger motions can be increased. Furthermore the virtual hand model would move more realistic, since the flexion of one finger depends only on one degree of motion. For example, the system could solely train the flexion of the distal phalanges (mostly flexor digitorum profundus) or the flexion of the middle phalanges (mostly flexor digitorum superficialis).

A further potential of improvement is the way of grabbing the ultrasound images. Instead of implementing a VGA grabber, the images obtained by the ultrasound system could be processed immediately and the information transferred to the machine learning unit. The question also arises if it is necessary at all to create an ultrasound image, or if it is possible to access and process the ultrasound system's information directly. This leads to completely new possibilities in terms of the setup specification and would decrease the computation time to a minimum.

List of Figures

4.1	Partial reflection of an ultrasound beam (adopted from Hedrick, 1995)	8
4.2	Series of echoes of an ultrasound beam (adopted from Hedrick, 1995)	9
4.3	Principle of echo ranging (adopted from Hedrick, 1995)	12
4.4	Main system components of a basic scanner (adopted from Hedrick, 1995)	13
4.5	Two applications of B-mode scanning: Prenatal screening of a foetus on the left and the visualisation of musculoskeletal structures (human forearm) on the right .	14
4.6	Timing pattern for a sequential linear array (adopted from Hedrick, 1995)	15
4.7	Position of cross section A and cross section B; right forearm (reproduced from Netter, 2006)	20
4.8	Cross section A: Anatomical structure of the right forearm; distal (reproduced from Netter, 2006)	21
4.9	Anatomical image compared to ultrasound image (cross section A)	21
4.10	Cross section B: Anatomical structure of the right forearm (mid-arm)	22
4.11	Anatomical image compared to ultrasound image (cross section B)	22
4.12	Bones of the human hand (adopted from Gray, 2008)	23
4.13	Parameters alpha, beta and gamma of a surface first order	32
5.1	Principal outline of the setup	35
5.2	Ultrasound system: <i>General Electric LOGIQ e</i> system	37
5.3	Ultrasonic transducer <i>12L-RS</i>	38
5.4	VGA to Ethernet frame grabber by <i>Epiphan Systems Inc.</i>	40
5.5	Tear frame detection	41
5.6	A close up of tear frame detection	42
5.7	Cyberglove with 18 sensors. Only the six sensors 1,4,5,7,10 and 13 are used . . .	43
5.8	Flexion of the pinkie demonstrated using one degree of freedom	45
5.9	Three phases of data gathering, training phase, and testing phase	45
5.10	Outline of the off-line setup	46
5.11	Outline of the on-line setup	47
5.12	Example of a grayscale ultrasound image	49
5.13	Grid points	50
5.14	Grid points with circle areas	50
5.15	Feature points with corresponding areas	51
5.16	Number of grid points change when adjusting the radius and point to point distance	51
5.17	Sequence of six image features	52
5.18	Interpolation of new finger position data	54
5.19	Smoothed ultrasound image before applying the optical flow algorithm	58
5.20	Vector components obtained after applying the optical flow algorithm	58
5.21	A close of the vector components obtained after applying the optical flow algorithm	59
5.22	Shifting feature points	59

6.1	Different finger movements including resting position	63
6.2	Setup used in preliminary experiments	64
7.1	Comparison of ultrasound images of different subjects	68
7.2	Comparison of ERR before (blue bars at back) and after (green bars at front) probe motion compensation. Bars denote the mean and standard deviation obtained on the 100 permutations.	72
7.3	Comparison of NRMSE before (purple bars at back) and after (orange bars at front) probe motion compensation. Bars denote the mean and standard deviation obtained on the 100 permutations.	73
7.4	Comparison of CC before (red bars in front) and after (brown bars at back) probe motion compensation. Bars denote the mean and standard deviation obtained on the 100 permutations.	74
7.5	Target values (black line) and predicted values with probe motion compensation (blue dotted line) and without (red dotted line)	74
7.6	Regression accuracy depends on training set size. The system was trained on every second (first row) every sixth (second row) or every tenth (third row) sample of the data set	75
7.7	Correlations between features and degrees of freedom	76
7.8	Fixation of the forearm	78
7.9	Experiment 1: The subject is training the system by mimicking the movements on the screen	79
7.10	Experiment 1: Testing phase	79
7.11	Results for all six degrees of motion. Target position values represented by the black line, predicted values by the red line	80
7.12	Experiment 2: The subject trains the system moving the imaginary limb coherently with the left hand	81
7.13	Testing phase of the second experiment	81
7.14	Results for all six degrees of motion. Target position values represented by the black line, predicted values by the red line	82
7.15	Correlation coefficient between image feature points and finger movements. The system was trained on the first half of the data set and tested on the other half	83
7.16	Mean absolute error for the system trained on the first half of the data set and tested on the other half	84
7.17	Correlation coefficient between image feature points and finger movements with shuffling the samples before training the system	84
7.18	Mean absolute error for the system. Samples were shuffled before training	85

List of Tables

4.1	Properties of different media (Hedrick, 1995)	7
4.2	Percentage reflection at different interfaces (Hedrick, 1995)	9
4.3	Attenuation of human tissues at 1MHz (Hedrick, 1995)	11
4.4	Terms of direction (Crouch, 1978)	17
4.5	Terms of movements (Crouch, 1978)	17
4.6	Muscles of the forearm that move the wrist and fingers (Montgomery, 1981) . . .	19
4.7	Interior muscles of the hand that move the fingers (Montgomery, 1981)	19
4.8	Commonly used treatments for phantom limb pain (Flor, 2002)	34
5.1	Optimised setting for the ultrasound system (see also General Electric Co., 2006b)	38
6.1	Different finger movements	62
7.1	Error results obtained by linear regression on each degree of motion: mean absolute error (ERR), normalised square-root mean-square-error (NRMSE) and correlation coefficient (CC).	69
7.2	Error results obtained after applying probe motion compensation: mean absolute error (ERR), normalised square-root mean-square-error (NRMSE) and correlation coefficient (CC).	70
7.3	ERR, NRMSE and CC for a system with various training set size (n denotes the step size)	71
7.4	Pre-experiment questionnaire	73
7.5	Pre-experiment questionnaire	75
7.6	Usage of prosthesis	77

List of Abbreviations

DIP	Distal Interphalangeal Joint
DLR	Deutsches Zentrum für Luft- und Raumfahrt
CC	Correlation Coefficient
EMG	Electromyography
ERR	Mean Absolute Error
MCP	Metacarpophalangeal Joint
ML	Machine Learning
NRMSE	Normalised Square-Root Mean-Square-Error
PIP	Proximal Interphalangeal Joint
SVM	Support Vector Machine
UDP	User Datagram Protocol
VGA	Video Graphics Array

Bibliography

- Bartsch, H.J., 2010. *Kleine Formelsammlung Mathematik*. 3rded. Munich, Vienna: Carl Hanser Verlag.
- Bergman, R.A., 2006. Illustrated Encyclopedia of Human Anatomic Variation, [Online] Available at: <<http://www.anatomyatlases.org/AnatomicVariants>> [Accessed 6 July 2011].
- Bishop, C.M., 2007. *Pattern recognition and machine learning*. New York: Springer.
- Bitzer, S. van der Smagt, P., 2006. Learning EMG control of a robotic hand: towards active prostheses. In: *IEEE International Conference on Robotics and Automation*. Orlando, USA May 2006.
- Burges, J.C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2 (2), pp.121-167.
- Castellini, C. van der Smagt. P., 2009. Surface EMG in advanced hand prosthetics. *Biological Cybernetics*, 100 (1), pp.35-47.
- Chen, X. Zheng, Y.P. Guo, J.Y. Shi, J., 2010. Sonomyography (SMG) control for powered prosthetic hand: a study with normal subjects. *Ultrasound in Med. & Biol*, 36 (7), pp.1076–1088.
- Crouch, J.E., 1978. *Functional Human Anatomy*. 3rd ed. Philadelphia: Lea & Febiger.
- De Flaviis L. et al., 1988. Ultrasonography of the hand in rheumatoid arthritis. *Acta Radiol*, 29, pp.457–460.
- Epiphan Systems Inc., 2009. *VGA2Ethernet User Guide 10.0.1*.
- Flor, H., 2002. Phantom-limb pain: characteristics, causes, and treatment. *Lancet Neurology*, 1, pp.182-189.
- Fuchs, M. et al. 2009. Rollin' Justin – Design considerations and realization of a mobile platform for a humanoid upper body. In: *International Conference on Robotics and Automation (ICRA)*. Kobe, Japan 12-17 May 2009.
- Gagné, C. et al., 2009. Motor control over the phantom limb in above-elbow amputees and its relationship with phantom limb pain. *Neuroscience*, 162, pp.78-86.
- General Electric Co., 2006. *LOGIQ e Basic User Manual*.
- General Electric Co., 2006. *LOGIQ e Service Manual*.
- Golub, H., 1996. *Matrix computations*. Baltimore: John Hopkins University Press.
- Gray, H., 2008. *Gray's anatomy*. Edinburgh: Churchill Livingstone.

- Grebenstein, M. et al. 2011. The DLR Hand Arm System. In: *IEEE International Conference on Robotics and Automation*. Shanghai, China 9-13 May 2011.
- Guo, J.Y. et al., 2011. A comparative evaluation of sonomyography, electromyography, force, and wrist angle in a discrete tracking task. *Ultrasound in Med. & Biol.*, 37 (6), pp.884–891.
- Hedrick, W.R., 1995. *Ultrasound Physics and Instrumentation*. St. Lois: Mosby.
- Hillmann, M. et al., 2004. Rehabilitation robotics from past to present - a historical perspective. *Lecture Notes in Control and Information Sciences*, 306, pp.25-44.
- Kandel, E.R. Schwartz, J.R. Jessell, T.M., 2000. *Principles of Neural Science*. New York [et al.]: McGraw-Hill.
- Liu, H. et al. 2008. Multisensory Five-Finger Dexterous Hand: The DLR/HIT Hand II. In: *IEEE International Conference on Intelligent Robots and Systems (RSJ)*. Nice, France 22-26 September 2008.
- Meier, R.H. and Atkins, D.J., 2004. *Functional Restoration of Adults and Children with Upper Extremity Amputation*. New York: Demos Medical Publishing.
- Montgomery, R.J., 1981. *Basic Anatomy for the Allied Health Professions*. 3rd ed. Baltimore, Munich: Urban & Schwarzenberg.
- MVTec Software GmbH, 2010. *HALCON Reference Manual 6.1*.
- Netter, H., 2006. *Atlas der Anatomie des Menschen*. 3rd ed. Stuttgart: Thieme.
- Nilsson, N.J., 1998. *Introduction to Machine Learning*. Stanford: Stanford University.
- Ramachandran, V.S. Rogers-Ramachandran, D., 1996. Synaesthesia in Phantom Limbs Induced with Mirrors. *Proc. R. Soc. Lond. B*, 263 (1), pp.377–386.
- Rao, C.R., 1971. *Generalised inverse of matrices and its applications*. New York: Wiley.
- Ripley, B.D., 2002. *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.
- Russ, J.C., 2006. *The image processing handbook*. 5th ed. Boca Raton: CRC Press.
- Sebastin, S.J. et al., 2005. Does the Absence of the Palmaris Longus Affect Grip and Pinch Strength?. *Journal of hand surgery (European volumes)*, 30 (4), pp.406–408.
- Shi, J. et al. 2010. Recognition of Finger Flexion from Ultrasound Image with Optical Flow: A Preliminary Study. In: *International Conference on Biomedical Engineering and Computer Science (EMBC)*. Wuhan, China 23-25 April 2010.
- Shih, F.Y., 2010. *Image processing and pattern recognition*. Piscataway: IEEE Press.
- Speich, J.E. and Rosen, J., 2004. Medical Robotics. *Encyclopedia of Biomaterials and Biomedical Engineering*. pp.983-993.
- Theodoridis, S. Pikrakis, A. Koutroumbas, K. Cavouras, D., 2010. *Introduction to Pattern Recognition. A MATLAB Approach*. Amsterdam: Elsevier.

- van der Smagt, P. et al., 2009. Robotics of human movements. *Journal of Physiology - Paris*, 103, pp.119–132.
- van der Smagt, P. Castellini, C. and Urbanek, H., 2009. Multisensory Five-Finger Dexterous Hand: The DLR/HIT Hand II. In: *International Conference on Multimodal Interfaces for Skills Transfer*. Bilbao, Spain 15-16 December 2009.
- World Health Organisation Study Group, 1998. Training in diagnostic ultrasound: essentials, principles and standards, [Online] Available at: <http://whqlibdoc.who.int/trs/WHO_TRS_875.pdf> [31 August 2011].
- Wirta, R.W. Taylor, D.R. Finley, F.R., 1977. *Pattern Recognition Arm Prosthesis: A historical perspective - A final report*. Philadelphia: Moss Rehabilitation Hospital.
- Zheng, Y.P. et al., 2006. Sonomyography: Monitoring morphological changes of forearm muscles in actions with the feasibility for the control of powered prosthesis". *Medical Engineering & Physics*, 28, pp.405–415.